

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:19:25 ; Search time 1773.53 Seconds
(without alignments)
13149.734 Million cell updates/sec

Title: US-09-802-937-2
Perfect score: 1440
Sequence: 1 ctgcaggaagattaagg.....cgcgagcgagtgctgac 1440

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_othr.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 99.4 | 6.9 | 637 | 17 | AQ510538 nbxb0095H |
| 2 | 99.4 | 6.9 | 813 | 17 | AQ287155 nbxb0026L |
| 3 | 98.4 | 6.8 | 627 | 17 | AQ257106 nbxb0017B |
| 4 | 95.8 | 6.7 | 488 | 17 | AQ867664 nbxb0033C |
| 5 | 95.8 | 6.7 | 718 | 17 | AQ859314 nbxb0012C |
| 6 | 95.2 | 6.6 | 757 | 17 | AQ290644 nbxb0037M |

| | | | | | |
|---|------|-----|------|----|----------|
| 7 | 94 | 6.5 | 1111 | 9 | AB021959 |
| c | 93.6 | 6.5 | 749 | 17 | AQ160799 |
| 8 | 90.4 | 6.3 | 1101 | 17 | CNS00EVL |
| c | 87.2 | 6.1 | 859 | 17 | AQ045395 |
| c | 86 | 6.0 | 746 | 17 | AQ160781 |
| c | 80.4 | 5.8 | 987 | 17 | CNS014PQ |
| c | 79 | 5.5 | 1101 | 17 | CNS0039G |
| c | 77 | 5.4 | 1101 | 17 | CNS00B07 |
| c | 77.8 | 5.4 | 1101 | 17 | CNS001PB |
| c | 77 | 5.3 | 1187 | 17 | B11102 |
| c | 76.6 | 5.3 | 1101 | 17 | CNS00PMC |
| c | 76.4 | 5.3 | 634 | 17 | AQ271541 |
| c | 76.4 | 5.3 | 730 | 17 | AQ289326 |
| c | 75.2 | 5.2 | 1101 | 17 | CNS003BD |
| c | 75 | 5.2 | 734 | 17 | AQ992421 |
| c | 72.8 | 5.1 | 543 | 17 | AQ156207 |
| c | 72.8 | 5.1 | 543 | 17 | AQ290984 |
| c | 72.8 | 5.1 | 556 | 17 | AQ288223 |
| c | 70.6 | 4.9 | 776 | 17 | AQ271490 |
| c | 69.6 | 4.8 | 1101 | 17 | CNS0039G |
| c | 69 | 4.8 | 1101 | 17 | CNS003B6 |
| c | 68.4 | 4.8 | 552 | 17 | AQ576900 |
| c | 68.4 | 4.8 | 1101 | 17 | CNS00EVL |
| c | 68.2 | 4.7 | 868 | 17 | AQ576154 |
| c | 68.2 | 4.7 | 1101 | 17 | CNS00DGI |
| c | 67.8 | 4.7 | 634 | 17 | AQ327958 |
| c | 67.4 | 4.7 | 1225 | 17 | CNS0161D |
| c | 67.2 | 4.7 | 994 | 17 | CNS04NOJ |
| c | 67.2 | 4.7 | 1101 | 17 | CNS00LT2 |
| c | 66.8 | 4.6 | 1101 | 17 | CNS00KAE |
| c | 66.6 | 4.6 | 346 | 17 | AQ954431 |
| c | 66.2 | 4.6 | 1025 | 17 | CNS014J2 |
| c | 66 | 4.6 | 1147 | 17 | B13042 |
| c | 65.8 | 4.6 | 1101 | 17 | CNS0021J |
| c | 65.6 | 4.6 | 609 | 17 | CNS025K2 |
| c | 65.6 | 4.6 | 765 | 17 | AQ315828 |
| c | 65.4 | 4.5 | 1037 | 14 | BQ48574 |
| c | 65.4 | 4.5 | 1092 | 17 | CNS020K7 |
| c | 65.4 | 4.5 | 1101 | 17 | CNS00EO7 |

ALIGNMENTS

RESULT 1
AQ510538
LOCUS nbxb0095H17r CUGI Rice BAC Library Oryza sativa genomic clone
DEFINITION 637 bp DNA linear GSS 04-MAY-1999
ACCESSION AQ510538
VERSION AQ510538.1 GI:4733142
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa

REFERENCE
AUTHORS Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 95
High quality sequence stop: 424.
Location/Qualifiers

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphorbia; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 637)

AB021959 AB021959
AQ160799 nbxb0006C
AL069706 Drosophil
AZ045395 nbxb0081C
AQ160781 nbxb0006A
AL104456 Drosophil
AL063921 Drosophil
AL069440 Drosophil
AL060732 Drosophil
B11102 F19C22-T7 I
AQ271541 Drosophil
AQ289326 nbxb0034L
AL064091 Drosophil
AQ992421 nbxb0091G
AQ156207 nbxb0002P
AQ290984 nbxb0038E
AQ288223 nbxb0032N
AQ271490 nbxb0025P
AL063921 Drosophil
AL069706 Drosophil
AQ576154 nbxb0088N
AL069971 Drosophil
AQ327958 nbxb0042I
AL106171 Drosophil
AL298972 Tetraodon
AL078714 Drosophil
AL077628 Drosophil
AQ954431 nbxb0070J
AL104216 Drosophil
B13042 T30M24-Sp6
AL061936 Drosophil
AL182171 Tetraodon
AQ915828 nbxb0060I
BQ48574 AGENCOURT
AL175696 Tetraodon
AL069440 Drosophil

```

source
1. .637
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbxb0095H17r"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT      207 a 134 c 121 g 173 t 2 others
ORIGIN

Query Match      6.9%; Score 99.4; DB 17; Length 637;
Best Local Similarity 74.6%; Pred. No. 6.8e-10;
Matches 138; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

Qy 893 TTTCAAACGCATGATAACGAGAAAGCTCATAGCACATTATTACTTAGATATTATTAAT 952
|||
Db 423 TTTCCAGCACACGCAAAACGAGAAACTCCTAGCACATGATTAAATTAAGTATTAACTAT 482

Qy 953 TATAAACTTGAAGAAAATATTTAT-TTCAATTTTAAACAATGTATGCATAAATTTAT 1011
|||
Db 483 TATAAAATTTGAAGAAAATAGATGTATTTAATTTATTAACAACCTCTATATAGAACTTT 542

Qy 1012 TTTAAACACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAA 1071
|||
Db 543 TGTCAATAACGCACAATTTAACGGTTTAGAAAGCATGCTAACGGAAACGAGGAAGTTAA 602

Qy 1072 AGATT 1076
|||
Db 603 AGTTT 607

RESULT 2
AQ287155      813 bp      DNA      linear      GSS 03-DEC-1998
LOCUS      nbxb0026L24r CUGI Rice BAC Library Oryza sativa genomic clone
DEFINITION      nbxb0026L24r, DNA sequence.
ACCESSION      AQ287155
VERSION      AQ287155.1 GI:3948897
KEYWORDS      GSS.
SOURCE      Oryza sativa.
ORGANISM      Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartodeae; Oryzeae; Oryza.
1 (bases 1 to 813)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Seq primer: GGAACACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 473.
Location/Qualifiers
1. .813
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbxb0026L24r"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT      259 a 182 c 138 g 234 t
ORIGIN

Query Match      6.9%; Score 99.4; DB 17; Length 813;
Best Local Similarity 74.6%; Pred. No. 6.5e-10;
Matches 138; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

Qy 893 TTTCAAACGCATGATAACGAGAAAGCTCATAGCACATTATTACTTAGATATTATTAAT 952
|||
Db 369 TTTCCAGCACACGCAAAACGAGAAACTCCTAGCACATGATTAAATTAAGTATTAACTAT 428

Qy 953 TATAAACTTGAAGAAAATA-TTTATTTGAATTTTAAACAATGTATGCATAAATTTAT 1011
|||
Db 429 TATAAAATTTGAAGAAAATAGATTTATTTAATTTATTAACAACCTCTATATAGAACTTT 488

Qy 1012 TTTAAACACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAA 1071
|||
Db 489 TGTCAATAACGCACAATTTAACGGTTTAGAAAGCATGCTAACGGAAACGAGGAAGTTAA 548

Qy 1072 AGATT 1076
|||
Db 549 AGTTT 553

RESULT 3
AQ257106      627 bp      DNA      linear      GSS 23-OCT-1998
LOCUS      nbxb0017B06r CUGI Rice BAC Library Oryza sativa genomic clone
DEFINITION      nbxb0017B06r, DNA sequence.
ACCESSION      AQ257106
VERSION      AQ257106.1 GI:3781588
KEYWORDS      GSS.
SOURCE      Oryza sativa.
ORGANISM      Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartodeae; Oryzeae; Oryza.
1 (bases 1 to 627)

```

AUTHORS Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence start: 4
 High quality sequence stop: 368.

FEATURES

source

1..627
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone_lib="nbx0017B06r"
 /clone="nbx0017B06r"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /notes="Vector: pBelBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
 202 a 135 c 116 g 173 t 1 others

Query Match 6.8%; Score 98.4; DB 17; Length 627;
 Best Local Similarity 74.1%; Pred. No. 1.1e-09;
 Matches 137; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
 Qy 893 TTTCAAACGCATGATAAACGAGAAAGCTCTATTAGCACATTATTACTTTAGATATTATAAT 952
 Db 387 TTTTCAGCACACGACGAAACGAGAAAGCTCTAGCACATGATTAATTAAGTATTACTAT 446
 Qy 953 TATAAACTTGAAAAAATA-TTTATTGAAATTTTAAACAATGTATGCAATAATTTT 1011
 Db 447 TATAAAATTTGAAAAAATAGATTTATTTAATTTTAAACAACCTCTATATAAAAACTNT 506
 Qy 1012 TTTTAAACACACCAATTTTAAACCTTTTAAACATCCTTAATAGGAACGAGCACTTAA 1071
 Db 507 TGTCAATACGCACAAATTTTAAACGTTTAGAAGCATCTTAACGAAAAACGAGGAAGTTTA 566
 Qy 1072 AGATT 1076
 Db 567 AGTTT 571

RESULT 4
 AQ867664 488 bp DNA linear GSS 03-NOV-1999
 LOCUS nbx0033C04f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
 DEFINITION clone nbx0033C04f, DNA sequence.
 ACCESSION AQ867664

VERSION AQ867664.1 GI:6218121
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa

REFERENCE 1 (bases 1 to 488)
 AUTHORS Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence start: 9
 High quality sequence stop: 461.

FEATURES

source

1..488
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbx0033C04f"
 /clone_lib="CUGI Rice BAC Library (EcoRI)"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /notes="Vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."
 177 a 74 c 83 g 154 t

BASE COUNT 177 a 74 c 83 g 154 t
 ORIGIN

Query Match 6.7%; Score 95.8; DB 17; Length 488;
 Best Local Similarity 73.4%; Pred. No. 3.7e-09;
 Matches 149; Conservative 0; Mismatches 52; Indels 2; Gaps 2;
 Qy 874 CCTCTAGTTGGAGGTGATTTTCAAACGCATGATAAACGAGAAAGCTCATTAGCACATTA 933
 Db 104 CATCCAGTTGGAGGTATATTTTC-AGCACATGCAAAACGAATGAACATCATTAGCATGA 162
 Qy 934 TTACTTAGATATTTATATTAATAACTTGAAAAAATATTTATTGAAATTTTTTAAACAA 993
 Db 163 TTAATTAATATTAATTAATAAATGAAAAATAGTATTATTATTTTTTAACTACT 222
 Qy 994 TGTATGCATAAATTTATTTTTTAAACCAACACCAATTTTAAACCTTTTAAAGCATCCTTAAT 1053
 Db 223 TATACATATAAACITTTTATCAAAAATACACC-ATTTAATAGTTTTAATAACATGCTAAC 281
 Qy 1054 AGGAAACGAGGAAGTTAAAGATT 1076

```

Db 282 CGAAACGAGGAAGTTGAAATTT 304
||||| ||||| ||||| ||||| |||||
RESULT 5
AQ859314 718 bp DNA linear GSS 03-NOV-1999
LOCUS nbeb0012C12r CUGI Rice BAC Library (ECORI) Oryza sativa genomic
DEFINITION clone nbeb0012C12r, DNA sequence.
ACCESSION AQ859314
VERSION nbeb0012C12r
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 718)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 15
High quality sequence stop: 440.
FEATURES
source
1..718
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbeb0012C12r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
haploid genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa,
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,296 clones with an average insert size
of 121 Kb providing approximately 15 haploid genome
equivalents. The deep coverage allows the isolation a
particular sequence with a probability of 99.9 %. Three
high density filters, each containing 18,432 clones
(doubly spotted), represent the whole library for colony
screening and can be requested from the Clemson University
BAC/EST Resource Center (www.genome.clemson.edu)."
BASE COUNT 247 a 127 c 131 g 211 t 2 others
ORIGIN
Query Match 6.78; Score 95.8; DB 17; Length 718;
Best Local Similarity 73.4; Pred. No. 3.4e-09;
Matches 149; Conservative 0; Mismatches 52; Indels 2; Gaps 2;
QY 874 CCTCTAGTTGGAGGTGATTTTCAACCGCATGATAAAGCGAAGAGCTCATTTAGCATTATTA 933
Db 151 CATCCAGTTGGAGGTATATTTC-AGCACATGCANAGATGAATCACTATTAGCACAATGA 209

```

```

QY 934 TTACTTAGATATTTATATTAATAAACTTGAAGAAATATTTATTTGAATTTTAAACAA 993
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 210 TTAATTAATATTAATTAATTAATAAAATGAAATAAGTAGATTATTTTAAACAAC 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 TGTATGCATAAATTTATTTTAAACACACACAAATTTAACCCCTTTAAAAAGCATCTAAT 1053
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 270 TATACATATAAACTTTTATCAAAAATACACC-ATTTAATAGTTTAAATAACATGCTAAC 328
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 AGGAACGAGGAAGTTAAGATT 1076
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 GGAACGAGGAAGTTGAAATTT 351
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
AQ290644 757 bp DNA linear GSS 03-DEC-1998
LOCUS nbxb0037M24f CUGI Rice BAC Library Oryza sativa genomic clone
DEFINITION nbxb0037M24f, DNA sequence.
ACCESSION AQ290644
VERSION AQ290644.1
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 757)
Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 421.
FEATURES
source
1..757
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbxb0037M24f"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."
BASE COUNT 243 a 165 c 131 g 217 t 1 others
ORIGIN

```


Query Match 6.6%; Score 95.2; DB 17; Length 757;
 Best Local Similarity 73.0%; Pred. No. 4.5e-09;
 Matches 135; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 893 TTTCACAGCATGATAACGAGAAAGCTATTAGCACATTTACTAGTATTTTATAT 952
 Db 370 TTTCACAGCACGCAACGAGAAAGCTACTAGCACATGATTAATTAAGTATTAACAT 429

QY 953 TATAAAGCTTGAATAAATA-TTTATTGTAATTTTTTAAACAATGTATGCATAAATTTAT 1011
 Db 430 TATAAATGGAATAAATAGATTTATTATTATTATTAGACAATCTTCTATAGAAACTTT 489

QY 1012 TTTAAACAACACCAATTTTAAACCTTTTAAAGATCCTTAATAGGAAACGAGGAAGTTAA 1071
 Db 490 TGTCAATAAGCGCAATNAACGTTTGTAGAAGCATGCTAACGGAACGAGGAAGTTAA 549

QY 1072 AGATT 1076
 Db 550 AGTTT 554

RESULT 7
 AB021959
 LOCUS AB021959 1111 bp mRNA linear EST 07-JAN-1999
 DEFINITION AB021959 Oryza sativa ms-bo-Taichung 65 immature anther Oryza
 sativa cDNA clone DD96-1-7, mRNA sequence.
 ACCESSION AB021959
 VERSION AB021959.1 GI:4107146
 KEYWORDS EST.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 1111)
 Toriyama, K. and Ohnashi, Y.
 EST of immature anther of rice
 Unpublished (1999)
 Contact: Toriyama K
 Faculty of Agriculture, Laboratory of Plant Breeding
 Tohoku University
 Tsutsumidori-Anamiyamachi 1-1, Aobaku, Sendai, Miyagi 981-8555,
 Japan
 Email: torikin@bios.tohoku.ac.jp.
 Location/Qualifiers
 1..1111
 /organism="Oryza sativa"
 /strain="ms-bo-Taichung 65"
 /db_xref="taxon:4530"
 /clone="DD96-1-7"
 /clone_lib="Oryza sativa ms-bo-Taichung 65 immature
 anther"
 /tissue_type="anther"
 /dev_stage="immature anther"
 335 a 184 c 265 g 327 t

BASE COUNT 335 a 184 c 265 g 327 t
 ORIGIN

Query Match 6.5%; Score 94; DB 9; Length 1111;
 Best Local Similarity 65.9%; Pred. No. 7.2e-09;
 Matches 153; Conservative 0; Mismatches 75; Indels 4; Gaps 1;

QY 849 AAACACAGCCTTAAGGCTTCTGTAGTCTCTAGTTGGAGGTTGATTTTCAACGCGATGATA 908
 Db 804 AAACATCCAGGCTGCGTTTCATCTACACTGTGCAGGTAGATTTTTTTAGTGATCATTA 863

QY 909 AACGAGAAAGCTCATTAGCACATTTACTTAGATATTTTATAATTATAAATTTGAAAAAA 968
 Db 864 AACGAGAAAGCTTATTAGCACATGATTAATTAAGTATTAATAATTATTTTAAAC 923

QY 969 ATATTTATTTGAATTTTAAACATGATGATGATAAATTTATTTTAAACACACCA-- 1026
 Db 924 AGATTTGTTGAATTTTAAACAACTTCTATATAGACTTTTTTTTTTAAAAAAAATACA 983

QY 1027 --ATTTAACCTTTAAAGCATCTCTAATAGGAACGAGGAAGTTAAAGATT 1076
 Db 984 TCATGTACAGCTTTAAAGATCGTGTATAGAAACCGAGAAAGTTGAAGTTT 1035

RESULT 8
 AQ160799/c
 LOCUS AQ160799 749 bp DNA linear GSS 12-SEP-1998
 DEFINITION nbxb0006C07f CUGI Rice BAC Library Oryza sativa genomic clone
 nbxb0006C07f, DNA sequence.
 ACCESSION AQ160799
 VERSION AQ160799.1 GI:3552496
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 749)
 Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATGACCTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 464.
 Location/Qualifiers
 1..749
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbxb0006C07f"
 /clone_lib="CUGI Rice BAC Library"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /note="vector: pBeloBAC11; Site 1: HindIII; Site 2:
 HindIII; Rice is one of two most popular grains in the
 world. Half of the world population especially those
 inhabiting highly populated areas of the humid tropics
 and subtropics, rely on rice as their primary source of
 carbohydrate. Monocotyledonous rice is a diploid plant
 (2n=24) with a haploid genome equivalent of 431 Mbp
 (Arumuganathan and Earle, 1991). The relatively small
 genome of rice, three times larger than that of
 Arabidopsis, makes it suitable for genomic studies. In
 order to facilitate positional cloning, physical mapping
 and genome sequencing of rice, we have constructed a BAC
 library from Oryza sativa, Nipponbare variety. The
 library contains 36,864 clones with an average insert size
 of 128.5 Kb providing 10.9 haploid genome equivalents. The
 deep coverage allows the isolation a particular sequence
 with a probability of 99.9 %. Two high density filters,
 each containing 18,432 clones (doubly spotted), represent
 the whole library for colony screening."
 252 a 145 c 114 g 232 t 6 others

BASE COUNT 252 a 145 c 114 g 232 t 6 others
 ORIGIN

Query Match 6.5%; Score 93.6; DB 17; Length 749;
 Best Local Similarity 68.8%; Pred. No. 9.4e-09;
 Matches 139; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 878 TAGTTGGAGGTTGATTTTCAACGCGATGATAACGAGAAAGCTCATTAGCATTTATTC 937
 Db 744 TAGTTTGNAGAGCTTTTTCAGNCCGCTGTGAATGAGANAGCCCATTCAGCATTTAA 685

QY 938 TTAGATATTTATTAATAAACTTGAAAAAATATTTTATTGAAATTTTTT---AAACAAT 994

```

684 TTAAGTATTAACTATTATANACTTGAATAATGGATTATTATTGATTTTAAAAAAGC 625
995 GTATGTCATTAATTTTAAAAACACACCAAAATTAACCCCTTAAAAAGCATCCTTAATA 1054
624 TTTTCTATAAAAGTTTTCGAAAAACCGCCCTTTAATAGTNTGAAAGGTGCTGATG 565
1055 GGAACGAGGAAGTTAAAGATT 1076
564 GAAACCAAGGAAGTTGATGTTT 543

RESULT 9
CNS000EVL                                1101 bp      DNA      linear      GSS 04-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL069706
VERSION     AL069706.1 GI:4949849
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster.
ORGANISM   Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mamoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
            p1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES             Location/Qualifiers
     source           1..1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /notes="end : T7"
BASE COUNT          419 a   91 c   60 g   299 t   232 others
ORIGIN

Query Match          6.3%; Score 90.4; DB 17; Length 1101;
Best Local Similarity 36.2%; Pred. No. 3.7e-08;
Matches 228; Conservative 107; Mismatches 285; Indels 10; Gaps 2;

574 ACAACCATGTTTCGTACTCACTCTAATTTGTAAATTTCTATTTCAGTCACAAAATTC 633
445 MMMHMAATYCTCAHTTWMWMMWMAATTTWAAWAAATTTATTAATWAAWAAWAW 504
634 CAATTTCAATTAAGAAAAATAAACGTAGACGGGTAAACCCACCACCATCTAAGGTTAG 693
505 WWAATTTTMMWWTWTTTWTAAWTTTAAWAAAAAAWAAATTAATTTAAWAAWATAWA 564
694 CGAGGGTGAAGTACGCAGCAAAATATGATGGTTTATATATATCA-----TTTTTT 745
565 TTAWAAATTTAAWAAWTTATATTAATTTWATAAATTTWATTATATAAAAAAATATTTT 624

```

```

QY 746 TTAATAACTTTCCATATAAATTTCTTTAGGAAACATATCATTTAATGGTTTGAATAACGT 805
DB 625 ATAAATTTTAAATTTTAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTT 684
QY 806 GCACATTAAGAAACTAAGACGATGAGTTGGGAAACAAGAGAAAAACACAGCCTTAAGGC 865
DB 685 AAATATAWAAAAAATAAAAAAATAWAAATTAATAAATTAATAAATAWAAATAWAAWAA 744
QY 866 TTCTTGATCTCTAGTTGGAGGTTGATTTTCAACGCGATGATAAATCGAGAAAGCTCATTA 925
DB 745 ATWAWATAATATWATAATATATATTTTAAWWTATWAAWWTATATAWATAWAAWAAW 804
QY 926 GCACATTATTACTTATAGATATTATTAATATAAATCTGAAATAAATATTATTGGAATTTT 985
DB 805 ATAAATATAWATAWAAWAAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 864
QY 986 TTAACCAATGATGCATAAATTA--TTTTTAAATAACACACCAATTTTAACCTTTAAAAA 1043
DB 865 AATWATAWAAAAAATAAATTAATTTTWTWTTTAAWATAAATAWATAWAAWAAAAA 924
QY 1044 GCATCCTAATAGGAAACGAGCAAGTTAAAGATTCCACCAAGTGTTCGGATTAATGAAAAAT 1103
DB 925 AAAAAATTAWAAWWTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 984
QY 1104 GGGGTGGATTGAATTTGGTAAATGAATCAGGTTAGGTTAAATATAATAAATGAAAGAG 1163
DB 985 ATWTATATATWTTAATWATAATATTTTAAWWTATATTTTAAWAAWTAATATATATATAW 1044
QY 1164 GGAGAAATGAATGGTTAGATGTTAAATGTT 1193
DB 1045 WWTAAATATAWAAWAAATTAATTTATATATAT 1074

RESULT 10
AZ045395/c                                859 bp      DNA      linear      GSS 08-MAR-2000
LOCUS      nbe0081C21f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
DEFINITION clone OSJNB0081C21f, DNA sequence.
ACCESSION  AZ045395
VERSION     AZ045395.1 GI:7207135
KEYWORDS   GSS.
SOURCE     Oryza sativa.
ORGANISM   Oryza sativa.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1 (bases 1 to 859)
REFERENCE   1
AUTHORS    Wing,R.A. and Dean,R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL    Unpublished (1998)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University,
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
Seq primer: GTAAACGACGCGCAGTG
Class: BAC ends
High quality sequence start: 132
High quality sequence stop: 324.
FEATURES             Location/Qualifiers
     source           1..859
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="OSJNB0081C21f"
                     /clone_lib="CUGI Rice BAC Library (EcoRI)"
                     /tissue_type="Leaf"
                     /lab_host="E. coli DH10B"
                     /notes="Vector: pBACindigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of

```


- Web : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaut at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelcBAC11.

KEYWORDS
SOURCE
ORGANIS

GSS.
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The_BDGP/Drosophila_melanogaster_BAC_library was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/Drosophila_BAC.htm.

| FEATURES | SOURCE |
|---|----------------------------|
| 1. Geographical Location: The study area is located in the northern part of the state, covering a total area of 1,200 square kilometers. | State Geographical Data |
| 2. Population: The population of the study area is approximately 1.5 million, with a significant portion being urban. | State Census Data |
| 3. Climate: The climate is semi-arid, with annual rainfall ranging from 400 to 600 millimeters. | State Meteorological Data |
| 4. Topography: The terrain is mostly flat, with some low hills and valleys. | State Topographical Data |
| 5. Vegetation: The vegetation is primarily dry forest and scrubland. | State Environmental Data |
| 6. Water Resources: The study area has several rivers and lakes, providing a significant source of water. | State Water Resources Data |
| 7. Infrastructure: The infrastructure is well-developed, with a network of roads and railways. | State Infrastructure Data |
| 8. Economy: The economy is primarily based on agriculture and services. | State Economic Data |
| 9. Healthcare: The healthcare system is well-developed, with a network of hospitals and clinics. | State Healthcare Data |
| 10. Education: The education system is well-developed, with a network of schools and universities. | State Education Data |

```

tname source
1 1001
/organism="Drosophila melanogaster"
/db xref="taxon:7227"
/clone="BACRO8K10"
/clone_lib="RPCI-98"
/notes="end : TEf3"

201 a 64 C 131 g 202 t 503 others
BASE COUNT
ORIGIN

```

Query Match 5.5%; Score 79; DB 17; Length 1101;
Best Local Similarity 19.9%; Pred. No. 6.9e-06;
Matches 128; Conservative 268; Mismatches 243; Gaps 2;
Indels 4;

| | | | |
|----|------|---|-----|
| Qy | 357 | AATGAATACAAATTCCGCAAAATATCATGGTTATCTATCTTCGTCCAATTTGAAATTTTGAGT | 416 |
| Db | 1042 | WWWWWATWDTWMDKWWWWATAAKTDTAWTWRTAWRADWAGRDGAKGRDRAADTAAD | 983 |
| Qy | 417 | CCAACTGAGACTCAAATACGATTTTTCTTTTCCAAAAGAATAATTATTAAATTTTTTTTTTCAT | 476 |
| Db | 982 | GAGRDRGRKKOKKOGDDDKKGKKKAACAUKWATKWMDWDNDKDWKDGAKOR | 923 |
| Qy | 477 | GAACCGCAATTC AACCGTTCGAGAAATATGCTGCTCAATAAATAGTAGTCTTAGTGCAGAA | 536 |
| Db | 922 | KADDGCGKDGDGKDADDTDGTGKDDDDKDKWDWKAKGTGWGDATWAWAAATDWWW | 863 |
| Qy | 537 | CAAAAATTAATACACATAAAAAAGAGTTGTTAAATTACAAACCATGTTTCGTACTACAA | 596 |
| Db | 862 | WGWAADAMWTWDAAADDDWADDWDAMWKWDADAANGARTADRRDWDGPRAGKRGARKR | 803 |
| Qy | 597 | CTCTAAATTCGTAATCTTATTTCAGTCACAA-AAATTCCAATTTCCAATTTAAGAAAAATA | 655 |
| Db | 802 | RDRKBDDKDAADDRDDAATWTTTTRDTDHKKWKTDITWRWAAORTWBDRDDDDR | 743 |
| Qy | 656 | AACGTAGACGGCTAAGCCCACCACCTCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCCAGAA | 715 |
| Db | 742 | DRAFTGKRWRRTWKRWKRRTGTWDDADADDTARDRRRRRGDDGADAGKCKTGTGRKR | 683 |
| Qy | 716 | AAATATCATCGTTTATTAAATATCATTTTTTTTTTAAATAAATCTTCCACATAAATTTCTTTAGG | 775 |
| Db | 682 | RDRATWDRDADAWADAANWTTTDTDDDDKDRRRRKGBARRRRTTARAADWHTWKAWD | 623 |
| Qy | 776 | AAACATATCATTTAATGGTTTGA AAAACCGTCACATAAGAAAACTAAGAACCGATGAGTTG | 835 |
| Db | 622 | WAKWDKTKRADRWDRWAADTWTDAR KADRDKWA KAARWARRRDRARAARADRRRTTGKTT | 563 |

| | | | | | |
|------------|--|---------|-----|--------|-----------------|
| RESULT 13 | CNS0039G | 1101 bp | DNA | linear | GSS 03-JUN-1999 |
| LOCUS | CNS0039G/c | | | | |
| DEFINITION | Drosophila melanogaster genome survey sequence TPT3 end of BAC # | | | | |
| | BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit | | | | |
| | fly), genomic survey sequence. | | | | |
| ACCESSION | AL063921 | | | | |
| VERSION | AL063921.1 | | | | GI:4941778 |

```

QY 836 GGAACAACAGAGAAAACACAGCCCTTAAGGCTTCCTGATCCTCTAGTGGAGGTTGATTTT 895
Db 562 TATTTTAAARAANAANAANAATATATTTTATTTTATTTTATTTTATTTTATTTTAAWAAWTA 503
QY 896 CAAACGCATGATAAAGCAGAAAGCTCATTAGCACATTTACTTACATTTAGATTTTATAATAT 955
Db 502 TWAA--WTAANAANAANAANAATTTTATTTTWTAAWAAWTAATTTTATTTTATTTTAA 446
QY 956 AAACCTGAAAAAATATTTATTTGAAATTTTAAACAATGTAT 998
Db 445 AATTTTWTWTAAATATTTTATTTTATTTTAAANAANAANAANAANAANAANAANAANAANA 403

RESULT 14
CNS00E07
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC;
BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069440.1 GI:4949583
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29P01"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 366 a 66 c 104 g 351 t 214 others
ORIGIN
Query Match 5.5%; Score 79; DB 17; Length 1101;
Best Local Similarity 37.3%; Pred. No. 6.9e-06;
Matches 176; Conservative 67; Mismatches 229; Indels 0; Gaps 0;

QY 718 ATATGATGTTTATTAATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 777
Db 534 AAATAATTTTWTAAATAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 593
QY 778 ACATATCATTTTATGTTTGAANAACGTCACATATAAATAAACAATGATGTTGGG 837
Db 594 TTTTATATATTAAAGWAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 653
QY 838 AAACAAGAGAAAAACACAGCCCTTCTTGATCCTCTAGTGGAGGTTGATTTTCA 897
Db 654 TAAATTTTATTAATTTAAWTTTAAWAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAW 713

```

```

QY 898 AACGCATGATAAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTTATAATTATAA 957
Db 714 TTTTAAATAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 773
QY 958 ACTTGAAAAAATATTTATTTGAAATTTTAAACAATCTATGCATAAAATTTTATTTTAAA 1017
Db 774 WTTTAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 833
QY 1018 AACACACCAATTTAACCCCTTTAAAAGCATCTTAATAGGAAACGAGGAGTTAAAGATTC 1077
Db 834 TTTTAAATTTTATTTTATTTTAAWTTTATTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTT 893
QY 1078 ACCGAAGTCTTTGGATATGAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 1137
Db 894 AAAAANAATGTTTAAATTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAAT 953
QY 1138 TAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1189
Db 954 TATGTATAWAAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 1005

RESULT 15
CNS001FB/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR04A23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL060732.1 GI:4939397
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR04A23"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 288 a 110 c 103 g 491 t 109 others
ORIGIN
Query Match 5.4%; Score 77.8; DB 17; Length 1101;
Best Local Similarity 38.3%; Pred. No. 1.2e-05;
Matches 215; Conservative 62; Mismatches 284; Indels 1; Gaps 1;

QY 433 TAGCATTTTCTTTTCAAAAAGAAATTAATTAATTTTCTTATGAACGCAATTAACCC 492

```

Db 1083 TATAATWTATATWAAAAATWAATAAAATATWAAATWATAWATAATAAAAAATATAAAATW 1024
Qy 493 GTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAAAATTAATATCACA 552
Db 1023 ATTAAAAATAAATWTAATAAAATWAAAWAAAAATTTWAATTAATTATATATNNAAAAAA 964
Qy 553 TAAAAAAGAGGTTGTAAATTACAAACCATGTTTCGTACTACAACTCTAAATTTGTAAAT 612
Db 963 AAWAAAAAAATTAATAWAAAAAATAAAWATAATAATATATAAAWWTAAATWTAATWN 904
Qy 613 CTTATTTCAGTCACAAAATTCCAAATTCCTCAATTAAGAAA-AAATAACGTAGACGGCTAAG 671
Db 903 ATTWAAAAATTAATAWAAATTAWAAAAATTAWAAAAATTAWAAAAAATWNTATAAAT 844
Qy 672 CCCACCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCACGAAATAATGATGTTTAT 731
Db 843 WAAAAATWNTWAAAAATWTTTTTAAATATATATATWTAATWTAATWAAAAAA 784
Qy 732 TAATATGATTTTTTTTAAATACTTTCACATAAAATTTCTTTAGGAAACATATCATTTAAT 791
Db 783 ATAWAAWATWNTWAAWNAWAAWAAATAATATAWNAWAAAAAATTWAWANATWMT 724
Qy 792 GGTTCGAAAACTGCACATAGAAAACTAAGAACGATGAGTTGGGAAACAAGAGAAAA 851
Db 723 WWTAAAAATAAAAAATAWATAAAATAWAAAAAATAAAAAAATAAAAAAATAAAAAA 664
Qy 852 CACAGCCTTAAGGCTTCTTGATCCTCTAGTTCGAGGTTGATTTTCAACGCATGATAAAC 911
Db 663 AAAAGNNGGGGGGGGGGGGGGGGHHYWTWHTWATNTNNTTATATATTA 604
Qy 912 GAGAAAGCTCATTAGCACATTATTACTTAGATATTATTAATTAATAAATTGAAAAATA 971
Db 603 TTATTACATTATTATTNTTATTACTATTATTATTATTATTATTATTATTATTATTATTA 544
Qy 972 TTTATTGAATTTTTTAAACAA 993
Db 543 TTAATACTATATTTCAACCCA 522

Search completed: March 15, 2003, 16:52:51
Job time : 1786.53 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:19:25 ; Search time 5346.47 Seconds
(without alignments)
13149.734 Million cell updates/sec

Title: US-09-802-937-1
Perfect score: 4341
Sequence: 1 ctcgaggaagattaattagg.....gtatatcttcattgttttt 4341

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estlin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pin.*
- 21: em_gss_vit.*
- 22: em_gss_fun.*
- 23: em_gss_mam.*
- 24: em_gss_mus.*
- 25: em_gss_other.*
- 26: em_gss_pro.*
- 27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|------------|-------|-------------|--------------|--------------------|
| Result No. | Score | Query Match | Length DB ID | Description |
| 1 | 448 | 10.3 | 680 17 | AQ257253 nbxb0017P |
| 2 | 439.6 | 10.1 | 520 17 | AQ289489 nbxb0035C |
| 3 | 238 | 5.5 | 502 17 | AQ915399 nbxb0056H |
| 4 | 173.2 | 4.0 | 287 9 | AU164649 AU164649 |
| 5 | 156.4 | 3.6 | 822 17 | BH704748 BOMOR39TR |
| 6 | 149.6 | 3.4 | 648 12 | BF473637 WHE030_F |

| | | | | | | |
|----|-------|-----|------|----|----------|----------|
| 7 | 149.6 | 3.4 | 680 | 10 | BE493591 | BE493591 |
| 8 | 148.6 | 3.4 | 514 | 10 | BE405815 | BE405815 |
| 9 | 148 | 3.4 | 774 | 14 | BQ840903 | BQ840903 |
| 10 | 147.6 | 3.4 | 596 | 10 | BE443438 | BE443438 |
| 11 | 146.4 | 3.4 | 612 | 14 | BQ243888 | BQ243888 |
| 12 | 146.4 | 3.4 | 719 | 14 | BQ246627 | BQ246627 |
| 13 | 145.2 | 3.3 | 864 | 12 | BG418644 | BG418644 |
| 14 | 144.8 | 3.3 | 593 | 10 | BE495096 | BE495096 |
| 15 | 144.8 | 3.3 | 639 | 10 | BE403629 | BE403629 |
| 16 | 143.2 | 3.3 | 420 | 9 | AJ486371 | AJ486371 |
| 17 | 143.2 | 3.3 | 596 | 14 | BQ753292 | BQ753292 |
| 18 | 143.2 | 3.3 | 599 | 13 | BM098903 | BM098903 |
| 19 | 143.2 | 3.3 | 600 | 10 | AV929110 | AV929110 |
| 20 | 143.2 | 3.3 | 603 | 10 | AV913445 | AV913445 |
| 21 | 143.2 | 3.3 | 634 | 10 | AV924023 | AV924023 |
| 22 | 143.2 | 3.3 | 644 | 10 | AV946226 | AV946226 |
| 23 | 143.2 | 3.3 | 648 | 14 | BQ458847 | BQ458847 |
| 24 | 143.2 | 3.3 | 655 | 13 | BJ463888 | BJ463888 |
| 25 | 143.2 | 3.3 | 679 | 10 | AV918748 | AV918748 |
| 26 | 143.2 | 3.3 | 742 | 13 | BJ468059 | BJ468059 |
| 27 | 143.2 | 3.3 | 746 | 10 | BE421072 | BE421072 |
| 28 | 143.2 | 3.3 | 906 | 12 | BG301342 | BG301342 |
| 29 | 142.2 | 3.3 | 560 | 13 | BJ467426 | BJ467426 |
| 30 | 141.6 | 3.3 | 518 | 13 | BJ459088 | BJ459088 |
| 31 | 141.6 | 3.3 | 531 | 13 | BJ451593 | BJ451593 |
| 32 | 141.6 | 3.3 | 811 | 12 | BF621279 | BF621279 |
| 33 | 138 | 3.2 | 404 | 14 | BM953435 | BM953435 |
| 34 | 137.2 | 3.2 | 482 | 14 | BQ236703 | BQ236703 |
| 35 | 136.4 | 3.1 | 564 | 14 | BQ163113 | BQ163113 |
| 36 | 136.4 | 3.1 | 585 | 14 | BQ528230 | BQ528230 |
| 37 | 136.4 | 3.1 | 582 | 9 | AI670171 | AI670171 |
| 38 | 136.4 | 3.1 | 583 | 12 | BF587853 | BF587853 |
| 39 | 136.4 | 3.1 | 602 | 14 | BM95705 | BM95705 |
| 40 | 136.4 | 3.1 | 613 | 14 | BQ294265 | BQ294265 |
| 41 | 136.4 | 3.1 | 657 | 9 | AI739940 | AI739940 |
| 42 | 136.4 | 3.1 | 1088 | 11 | AY109686 | AY109686 |
| 43 | 136.2 | 3.1 | 445 | 14 | BQ239883 | BQ239883 |
| 44 | 135.2 | 3.1 | 508 | 17 | BH693012 | BH693012 |
| 45 | 134 | 3.1 | 393 | 13 | BJ462915 | BJ462915 |

ALIGNMENTS

RESULT 1
AQ257253
LOCUS nbxb0017P08r CUGI Rice BAC Library Oryza sativa genomic clone 680 bp DNA linear GSS 23-OCT-1998
DEFINITION nbxb0017P08r, DNA sequence.
ACCESSION AQ257253
VERSION AQ257253.1 GI:3781735
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 680)
AUTHORS Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 47
High quality sequence stop: 390.
Location/Qualifiers

FEATURES

source
 1. .680
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbx0017P08r"
 /clone_lib="CUGI Rice BAC Library"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
 BASE COUNT 193 a 116 c 127 g 237 t 7 others
 ORIGIN

Query Match 10.3%; Score 448; DB 17; Length 680;
 Best Local Similarity 93.9%; Pred. No. 4.7e-83;
 Matches 477; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
 QY 2734 AGCTTGAATATGATG-TCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAA 2792
 Db |||||
 QY 2793 AAGGTAAGACATACAAATTTTGATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT 2852
 Db |||||
 QY 101 AAGGTAAGACATACAAATTTTGATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT 160
 Db |||||
 QY 2853 GCGGAAAGGGAACACAGTCACCGCTGATTAAGATGAATTTTCTGTGCCATTTAGCCA 2912
 Db |||||
 QY 161 GCGGAAAGGGAACACAGTCACCGCTGATTAAGATGAATTTTCTGTGCCATTTAGCCA 220
 Db |||||
 QY 2913 CTGGTGATATGTTGAGGCTGAGTGGCTGCTAAACTCCACTGGGATTAAGGCTAAG 2972
 Db |||||
 QY 221 CTGGTGATATGTTGAGGCTGAGTGGCTGCTAAACTCCACTGGGATTAAGGCTAAG 280
 Db |||||
 QY 2973 AAGCTATGCAAGGTAGTGTTTTAAAGAAACATATAGCAACAGAAATTAACACGAGGA 3032
 Db |||||
 QY 281 AAGCTATGCAAGGTAGTGTTTTAAAGAAACATATAGCAACAGAAATTAACACGAGGA 340
 Db |||||
 QY 3033 ATGGGTTTCTGATCTTTTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 3092
 Db |||||
 QY 341 ATGGGTTTCTGATCTTTTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 400
 Db |||||
 QY 3093 ACTTGGTTGTTGGGATTTATGATGAAGCATCAAGAAACTTCATGCCAGAAAGCTTTTA 3152
 Db |||||
 QY 401 ACTTGGTTGTTGGGATTTATGATGAAGCATCAAGAAACTTCATGCCAGAAAGCTTTTA 460
 Db |||||
 QY 3153 TCCTTGATGGTTTCCCTAGAACTGTGTTTCAAGACACAGAAAGGTGAGGTCCTTGGTCAATA 3212
 Db |||||
 QY 461 TCCTTGATGGTTTCCCTAGAACTGTGTTTCAAGACACAGAAAGGTGAGGTCCTTGGTCAATA 520
 Db |||||
 QY 3213 TGCACCGCTATATATAAGAGCTCCCTTTT 3240
 Db |||||
 QY 521 TGCACCTACTATATATACCTTGATCTT 548

RESULT 2

AQ289489 520 bp DNA linear GSS 03-DRC-1998
 LOCUS nbx0035C14f CUGI Rice BAC Library Oryza sativa genomic clone
 DEFINITION nbx0035C14f, DNA sequence.
 ACCESSION AQ289489
 VERSION AQ289489.1 GI:3950935
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1 (bases 1 to 520)
 AUTHORS Wing, R.A. and Dean, R.A.
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 348.
 FEATURES
 Location/Qualifiers
 1..520
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbx0035C14f"
 /clone_lib="CUGI Rice BAC Library"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
 BASE COUNT 148 a 91 c 108 g 173 t
 ORIGIN
 Query Match 10.1%; Score 439.6; DB 17; Length 520;
 Best Local Similarity 94.9%; Pred. No. 2.7e-81;
 Matches 465; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
 QY 2751 TCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAAAGGTAAGACATACAATT 2810
 Db |||||
 QY 2811 TTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCTCGGAAAGGGAACACAGT 2870
 Db |||||
 QY 66 TTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCTCGGAAAGGGAACACAGT 125
 Db |||||
 QY 2871 CACCGCTGATTAAGGATGAATTTTGTGTTGTCATTTTAGCCACTGCTGATATGTTGAGG 2930
 Db |||||
 QY 126 CACCGCTGATTAAGGATGAATTTTGTGTTGTCATTTTAGCCACTGCTGATATGTTGAGG 185

QY 2931 CTGAGTGGCTGCTAAACCTCCACTGGATTAAGCTTAAGCAAGCTATGACAGGTAG 2990
 |||||
 Db 186 CTGAGTGGCTGCTAAACCTCCACTGGATTAAGCTTAAGCAAGCTATGACAGGTAG 245
 |||||
 QY 2991 TTTTAAAGAAACATATAGCAACAGAAATTATACACGACGAGGAATGGTTCCTGATCTT 3050
 |||||
 Db 246 -TTTAAAGAAACATATAGCAACAGAAATTATACACGACGAGGAATGGTTCCTGATCTT 304
 |||||
 QY 3051 TTGTTCTTCTTCTTCTTCTAGGAGAGCTGTTTCTGATGACTGCTGTTGGGATTA 3110
 |||||
 Db 305 TTGTTCTTCTTCTTCTTCTAGGAGAGCTGTTTCTGATGACTGCTGTTGGGATTA 364
 |||||
 QY 3111 TTGATGAGCCATGAAGAAACTTCATGCCAGAAAGTTTATCTCTGATGGTTTCCCTA 3170
 |||||
 Db 365 TTGATGAGCCATGAAGAAACTTCATGCCAGAAAGTTTATCTCTGATGGTTTCCCTA 424
 |||||
 QY 3171 GAACTGTTGTTCAAGCACAGAGTGAGTCCCTGCTCAATATGACCGCTATATAAAG 3230
 |||||
 Db 425 AAATCGAGGTCAATCTCATATATGAGGCTCTTCCTTCCTCCGCTATATTTAAA 484
 |||||
 QY 3231 AGCTCCTTTT 3240
 |||||
 Db 485 TGTCTCTCTT 494
 |||||

RESULT 3
 A0915399/c
 LOCUS A0915399 502 bp DNA linear GSS 02-DEC-1999
 DEFINITION nbe0056H06r CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
 clone nbe0056H06r, DNA sequence.
 ACCESSION A0915399
 VERSION A0915399.1 GI:6511915
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 502)
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
 JOURNAL Unpublished (1998)
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence start: 28
 High quality sequence stop: 200.
 Location/Qualifiers
 1. 502
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbe0056H06r"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /notes="Vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;
 Rice is the most important food crop in the world. Half of
 the world population, especially those inhabiting highly
 populated areas of the humid tropics and subtropics, rely
 on rice as their primary source of carbohydrate.
 Monocotyledonous rice is a diploid plant (2n=24) with a
 haploid genome equivalent of 431 Mbp (Arumuganathan and
 Earle, 1991). The relatively small genome of rice, three
 times larger than that of Arabidopsis, makes it suitable

for genomic studies. In order to facilitate positional
 cloning, physical mapping and genome sequencing of rice,
 we have constructed a BAC library from Oryza sativa,
 Nipponbare variety using EcoRI as the cloning enzyme. The
 library contains 55,296 clones with an average insert size
 of 121 Kb providing approximately 15 haploid genome
 equivalents. The deep coverage allows the isolation a
 particular sequence with a probability of 99.9 %. Three
 high density filters, each containing 18,432 clones
 (doubly spotted), represent the whole library for colony
 screening and can be requested from the Clemson University
 BAC/EST Resource Center (www.genome.clemson.edu)."

BASE COUNT 182 a 97 c 99 g 123 t 1 others
 ORIGIN

Query Match 5.5%; Score 238; DB 17; Length 502;
 Best Local Similarity 100.0%; Pred. No. 3.3e-39;
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4104 ATTGCTTTGCTTTTTCACAAAGTTTACTGGAGAACCTTAATTCAAAGGAAGATGACAC 4163
 |||||
 Db 501 ATTGCTTTGCTTTTTCACAAAGTTTACTGGAGAACCTTAATTCAAAGGAAGATGACAC 442
 |||||
 QY 4164 AGCTGCAGTATTGAAGTCAAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTT 4223
 |||||
 Db 441 AGCTGCAGTATTGAAGTCAAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTT 382
 |||||
 QY 4224 CCTTAGCACTACGTTTTTAAATATTCAGATATTCCTTTTAGGATGTAGTCGACTTCAG 4283
 |||||
 Db 381 CCTTAGCACTACGTTTTTAAATATTCAGATATTCCTTTTAGGATGTAGTCGACTTCAG 322
 |||||
 QY 4284 TTAAGCGGATTCCTTCAGTTGCATTACAGTGTTCCTGTATATCTTCATTGTTTTT 4341
 |||||
 Db 321 TTAAGCGGATTCCTTCAGTTGCATTACAGTGTTCCTGTATATCTTCATTGTTTTT 264
 |||||

RESULT 4
 AUI64649
 LOCUS AUI64649 287 bp mRNA linear EST 03-APR-2002
 DEFINITION AUI64649 Rice root Oryza sativa (japonica cultivar-group) CDNA
 clone R0936, mRNA sequence.
 ACCESSION AUI64649
 VERSION AUI64649.1 GI:11173174
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group).
 ORGANISM Oryza sativa (japonica cultivar-group).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 287)
 TITLE Rice cDNA from root (2000)
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
 305-8602, Japan
 Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 PROJECT="RGP".
 Location/Qualifiers
 1. 287
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="R0936"
 /clone_lib="Rice root"
 /note="Prepared from seedling root." 2 others

BASE COUNT 91 a 52 c 60 g 82 t 2 others
 ORIGIN

Query Match 4.0%; Score 173.2; DB 9; Length 287;

```

Best Local Similarity 98.3%; Pred. No. 1.le-25;
Matches 175; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3310 GCAGCTTGATGAATGTTGGCCAAACAAAGGTACTAAAGTTGACAGGTTCTAAATTTTGC 3369
Db 86 GAAGCTTGATGAATGTTGGCCAAACAAAGGTACTAAAGTTGACAGGTTCTAAATTTTGC 145

QY 3370 AATTGATGATCAATGATGGAAGAACGAATATACCGGTCTGTTGGATCCACCCTCAAGTTGG 3429
Db 146 AATTGATGATCAATGATGGAAGAACGAATATACCGGTCTGTTGGATCCACCCTCAAGTTGG 205

QY 3430 TAGATCTTATCATACAAAATTTGCTCTCTCTAGACATCTCGGATCGATGATGAAGT 3487
Db 206 TAGATCTTATCATACAAAATTTGCTCTCTCTAGACATCTCGGATCGATGATGAAGT 263

RESULT 5
BH704748
LOCUS BH704748
DEFINITION BH704748.1 GI:18786587
VERSION BH704748
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 822)
AUTHORS Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOMOR39TF
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..822
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOMOR39"
/clone_lib="BO 2.3 KB"
/notes="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 193 a 178 g 172 g 279 t
ORIGIN

Query Match 3.6%; Score 156.4; DB 17; Length 822;
Best Local Similarity 64.4%; Pred. No. 3.6e-22;
Matches 282; Conservative 0; Mismatches 126; Indels 30; Gaps 2;

QY 3051 TTGTTCTCTTCTTATCTCTAGGAGAGCTTGTTCGTGATGACTGGTTGGGATTA 3110
Db 103 TTTTGTAAATATGTTTACAGGAGCGGTGTGTCATGAAGTGGTTGTGGTATAA 162

QY 3111 TTGATGAAGCCATGAAGAAACCTTCATGCCAGAAAGTGTATTCCTTGATGTTTCCCTTA 3170
Db 163 TTGATGAAGCCATGAAGAAACCTTCATGCCAGAAAGTGTATTCCTTGATGTTTCCCTTA 222

QY 3171 GAACGTGTTTCAAGACAGAGAGTGGTCCCTTGGTCAATATGCACCGCTATATAAAG 3230
Db 223 GAACGTGTTTCAAGACAGAGAGTGGTCCCTTGGTCAATATGCACCGCTATATAAAG 269

QY 3231 AGCTCCCTTTTGTATTAGAGCTGCTCTATATAAATGACAGAGTTTCTATCATTTGATCACT 3290

```

```

Db 270 CGTCGTCACTGTTTCTTCTGTCTCTCTTCTTCTTGTCTTATGCTTGTCTTTGAGTTGT 329
QY 3291 TTTCCTACTAAAAATGGTGACGCTTGATGAATGTTGGCCAAACAAAGGTACTAAGATTG 3350
Db 330 T-----TGCAGCTAGACGAGATGCTTAAAGAGCGGGGAACCTGGAATCG 372
QY 3351 ACAAGGTTCTAAATTTTGGCAATTTGCAATTTGATGATGCAATCTGGAAGAACGAATTTACCGTCTGT 3410
Db 373 ACAAGGTTCTCAACTTTGGCATCGATGACTCACTGCTGGAGGAAAGAAATCACCGGAGAT 432
QY 3411 GGATCCACCCTCAAGTGGTAGATCTTATCATACAAAATTTGCTCTCTAAGACTCTCTG 3470
Db 433 GGATCCACCCTCAAGTGGTAGATCTTATCATACAAAATTTGCTCTCTAAGACTCTCTG 492
QY 3471 GACTTGATGATGAAGTC 3488
Db 493 GAGTTGATGATGAAGAC 510

RESULT 6
BF473637
LOCUS BF473637
DEFINITION BF473637.1 GI:11542819
VERSION BF473637
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 648)
AUTHORS Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,
Seaton, C.L. and Tong, J.C.
TITLE The structure and function of the expressed portion of the wheat
genomes - 5-15 DAP spike cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 51055555773
Fax: 51055555818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
source
1..648
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0930_F04_K08"
/clone_lib="Wheat 5-15 DAP spike cDNA library"
/tissue_type="Spike"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse. Spikes at 5, 10 and 15 DAP were harvested,
total RNA and poly(A) RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised to
give pBluescript phagemids in the TJ Close lab (Choi,
Close, Fenton) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other authors
)."
BASE COUNT 173 a 147 c 169 g 159 t
ORIGIN

Query Match 3.4%; Score 149.6; DB 12; Length 648;

```

```

Best Local Similarity 91.9%; Pred. No. 9.5e-21;
Matches 158; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371
|||||
Db 441 AGCTGATGAATGTCGCCAAACAAGGTCTAAGGTTGACAAGGTTCTAAATTTGCAA 500
|||||

QY 3372 TTGATGATGCAATGTCGGAAGACGAAATACCGTTCGTTGGATCCACCATCAAGTGTA 3431
|||||
Db 501 TTGATGATGCAATGTCGGAAGACGAAATACTGCGCGTTGGATACACCATCGAGTGTA 560
|||||

QY 3432 GATCTTATCATACAAAATTTGCTCCTCAAGACTCCTCGACTTGATGATGT 3483
|||||
Db 561 GATCTTACCATACAAAATTTGCTCCTCAAGACTCCAGGAGTTGATGATGT 612
|||||

RESULT 7
BE493591 BE493591 680 bp mRNA linear EST 16-APR-2001
LOCUS WHE0568_B05_C10ZE Triticum monococcum vegetative apex cDNA library
DEFINITION Triticum monococcum cDNA clone WHE0568_B05_C10, mRNA sequence.
ACCESSION BE493591
VERSION BE493591.1 GI:9660184
KEYWORDS EST.
SOURCE Triticum monococcum.
ORGANISM Triticum monococcum.

REFERENCE
AUTHORS Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
Stamove,B. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat
genomes - vegetative apex cDNA library from Triticum monococcum
Unpublished (2001)
JOURNAL US Department of Agriculture, Agriculture Research Service, Pacific
COMMENT West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene T3 primer.
Location/Qualifiers
1. .680
/organism="Triticum monococcum"
/cultivar="DV92"
/db_xref="taxon:4568"
/clone_lib="WHE0568_B05_C10"
/library="Triticum monococcum vegetative apex cDNA
library"
/tissue_type="Vegetative shoot apex"
/dev_stage="Three weeks-old plants"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda pbk-CMV (Lambda Zap Express),
excised phagemid; Site 1: EcoRI; Site 2: XhoI; The tissue,
total RNA, and poly(A)+RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised at the
University of California, Davis (V. Echenique, B. Stamova
, J. Dubcovsky). Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
```

```

BASE COUNT 202 a 134 c 154 g 190 t
ORIGIN
```

```

Query Match 3.4%; Score 149.6; DB 10; Length 680;
Best Local Similarity 91.9%; Pred. No. 9.5e-21;
Matches 158; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371
|||||
```

```

Db 160 AGCTCGATGAATGTCGCCAAACAAGGTGCTAAGGTTGACAAGGTTCTAAATTTGCAA 219
|||||
QY 3372 TTGATGATGCAATGTCGCCAAACAAGATTACCGTTCGTTGGATCCACCATCAAGTGTA 3431
|||||
Db 220 TTGATGATGCAATGTCGCCAAACAAGATTACTGCGCGTTGGATACACCATCGAGTGTA 279
|||||

QY 3432 GATCTTATCATACAAAATTTGCTCCTCAAGACTCCTCGACTTGATGATGT 3483
|||||
Db 280 GATCTTACCATACAAAATTTGCTCCTCAAGACTCCAGGAGTTGATGATGT 331
|||||

RESULT 8
BE405815 BE405815 514 bp mRNA linear EST 21-JUL-2000
LOCUS WHE0437_C12_E23ZS Wheat etiolated seedling root cDNA library
DEFINITION Triticum aestivum cDNA clone WHE0437_C12_E23, mRNA sequence.
ACCESSION BE405815
VERSION BE405815.1 GI:9365283
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 514)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
TITLE The structure and function of the expressed portion of the wheat
genomes
Unpublished (2000)
JOURNAL Contact: Olin Anderson
COMMENT US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers
1. .514
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone_lib="WHE0437_C12_E23"
/clone_lib="Wheat etiolated seedling root cDNA library"
/tissue_type="Root"
/lab_host="E. coli SOLR"
/dev_stage="Five day old etiolated seedling"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Roots were
harvested. The tissue, total RNA, and poly(A)+ RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give Bluescript phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
```

```

BASE COUNT 136 a 115 c 128 g 134 t 1 others
ORIGIN
```

```

Query Match 3.4%; Score 148.6; DB 10; Length 514;
Best Local Similarity 91.8%; Pred. No. 1.5e-20;
Matches 157; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371
|||||
```

```

Db 344 AGCTGATCAAAATGCTGGCAAGCAAGGTGCTTAAGTTTGACAAGGTTCTAAATTTTGCAA 403
QY 3372 TTGATGATGCATATCTGGAGAACCAATACCGTCTGTGGATCCACCCATCAAGTGGA 3431
Db 404 TTGATGATGCATATCTGGAGAACCAATACCGTCTGTGGATCCACCCATCAAGTGGA 463
QY 3432 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGTGATGATG 3482
Db 464 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGTGATGATG 514

RESULT 9
BQ840903
LOCUS BQ840903 774 bp mRNA linear EST 12-AUG-2002
DEFINITION WHE4204_G10_N20Zs Aegilops speltoides pre-meiotic anther cDNA
library Aegilops speltoides cDNA clone WHE4204_G10_N20, mRNA
sequence.
ACCESSION BQ840903
VERSION BQ840903.1 GI:22210312
KEYWORDS EST:
SOURCE Aegilops speltoides.
ORGANISM Aegilops speltoides.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Aegilops.
REFERENCE 1 (bases 1 to 774)
AUTHORS Akhunov,E., Anderson,O.D., Chao,S., Crossman,C., Devorak,J., Lazo
,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
TITLE The structure and function of the expressed portion of the wheat
genomes - Pre-meiotic anther cDNA library from Aegilops speltoides
Unpublished (2002)
JOURNAL Contact: Olin Anderson
COMMENT US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
source
1..774
/organism="Aegilops speltoides"
/cultivar="P2 from 2-12-4-8-1-1-1-1" x P136909-12-811-1(1)
/db_xref="taxon:4573"
/clone="WHE4204_G10_N20"
/clone_lib="Aegilops speltoides pre-meiotic anther cDNA
library"
/tissue_type="Anther"
/dev_stage="Pre-meiotic anthers"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in a growth
chamber at the University of California, Davis (Akhunov in
Devorak's lab). Pre-meiotic anthers were harvested, total
RNA and poly(A) RNA were prepared, a cDNA library was made
, and the cDNA clones were in vivo excised to give
pBluescript phagemids at the University of California,
Davis (Akhunov in Devorak's lab). Plasmid DNA preparations
and DNA sequencing were performed in the OD Anderson lab
(all other authors)."
BASE COUNT 209 a 176 c 204 g 185 t
ORIGIN
Query Match 3.4%; Score 148; DB 14; Length 774;
Best Local Similarity 91.3%; Pred. No. 2.1e-20;
Matches 157; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3312 ACCTTGATGAATGTGGCCAAACAGGTACTAGATTGACAGGTTCTAAATTTTGCAA 3371
Db 467 AGCTGATGAATGTGGCCAAACAGGTGCTTAAGTTGACAGGTTCTAAATTTTGCAA 526

```

```

QY 3372 TTGATGATGCATATCTGGAGAACCAATACCGTCTGTGGATCCACCCATCAAGTGGA 3431
Db 527 TTGATGATGCATATCTGGAGAACCAATACCGTCTGTGGATCCACCCATCAAGTGGA 586
QY 3432 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGTGATGATG 3483
Db 587 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGAGTTGATGATG 638

RESULT 10
BQ443438
LOCUS BQ443438 596 bp mRNA linear EST 25-JUL-2000
DEFINITION WHE1104_B09_D18Zs Wheat etiolated seedling root normalized cDNA
library Triticum aestivum cDNA clone WHE1104_B09_D18, mRNA
sequence.
ACCESSION BQ443438
VERSION BQ443438.1 GI:9442972
KEYWORDS EST:
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 596)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,
Rausch,C.J., Seaton,C.L., Tong,J.C. and Zhang,B.
TITLE The structure and function of the expressed portion of the wheat
genomes - Normalized root cDNA library
Unpublished (2000)
JOURNAL Contact: Olin Anderson
COMMENT US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
source
1..596
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1104_B09_D18"
/clone_lib="Wheat etiolated seedling root normalized cDNA
library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli DH10B"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK; Site 1: EcoRI; Site 2: XhoI; Seeds were
surface-sterilized, germinated and grown aseptically in
the dark at room temperature on filter paper with water,
nystatin and cefotaxime in covered crystallization
dishes. Roots were harvested. The tissue, total RNA, and
poly(A) RNA were prepared, a cDNA library was made in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. The cDNA clones were in vivo
excised to give pBluescript phagemids before
normalization was carried out. The mass excision of
phagemid library and normalization were done in HT Nguyen
lab by D. Zhang at Texas Tech University. Normalization
protocol used was that of Soares. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 157 a 138 c 156 g 145 t
ORIGIN
Query Match 3.4%; Score 147.6; DB 10; Length 596;
Best Local Similarity 91.8%; Pred. No. 2.5e-20;

```

```

Matches 156; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAAGTTCTAAATTTTGCAA 3371
Db 427 AGCTCGATGAATGCTGGCAAGAGGTTGCTAAGGTTGACAAGGTTCTAAATTTTGCAA 486
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTTACCGGTCTGTTGGATCCACCCATCAAGTGGTA 3431
Db 487 TTGATGATGCAATCTGGAAGAACGAATTTACTGGCGTTGGATACACCCATCGAGTGGTA 546
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGAT 3481
Db 547 GATCTTACCATACAAAATTTGCTCTCTTAAGACTCCAGGAGTTGATGAT 596

RESULT 11
LOCUS BQ243888/c 612 bp mRNA linear EST 03-MAY-2002
DEFINITION TAE15007G03F TAE15 Triticum aestivum cDNA clone TAE15007G03F, mRNA
sequence.
ACCESSION BQ243888
VERSION BQ243888.1 GI:20439764
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 612)
AUTHORS Cloutier,S.
JOURNAL Wheat functional genomics - Glenlea developing seeds cDNA libraries
COMMENT Unpublished (2002)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 007 row: G column: 03
Seq primer: M13 Forward.
Location/Qualifiers
1..612
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03F"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"

BASE COUNT 180 a 131 c 125 g 176 t
ORIGIN

Query Match 3.4%; Score 146.4; DB 14; Length 612;
Best Local Similarity 90.7%; Pred. No. 4.4e-20;
Matches 156; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAAGTTCCTAAATTTTGCAA 3371
Db 555 AGCTCGATGAATGCTGGCAAGAGGTTGCTAAGGTTGACAAGGTTCTAAATTTTGCAA 496
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTTACCGGTCTGTTGGATCCACCCATCAAGTGGTA 3431
Db 495 TTGATGATGCAATCTGGAAGAACGAATTTACTGGCGTTGGATACACCCATCGAGTGGTA 436
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGAT 3483

```

```

Db 435 GATCTTACCATACAAAATTTGCTCTCTCCCAAGACTCCGGGAGTTGATGATGT 384

RESULT 12
LOCUS BQ246627 719 bp mRNA linear EST 03-MAY-2002
DEFINITION TAE15007G03R TAE15 Triticum aestivum cDNA clone TAE15007G03R, mRNA
sequence.
ACCESSION BQ246627
VERSION BQ246627.1 GI:204442503
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 719)
AUTHORS Cloutier,S.
JOURNAL Wheat functional genomics - Glenlea developing seeds cDNA libraries
COMMENT Unpublished (2002)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 007 row: G column: 03
Seq primer: M13 Reverse.
Location/Qualifiers
1..719
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03R"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"

BASE COUNT 207 a 148 c 163 g 201 t
ORIGIN

Query Match 3.4%; Score 146.4; DB 14; Length 719;
Best Local Similarity 90.7%; Pred. No. 4.4e-20;
Matches 156; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAAGTTCCTAAATTTTGCAA 3371
Db 230 AGCTCGATGAATGCTGGCAAGAGGTTGCTAAGGTTGACAAGGTTCTAAATTTTGCAA 289
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTTACCGGTCTGTTGGATCCACCCATCAAGTGGTA 3431
Db 290 TTGATGATGCAATCTGGAAGAACGAATTTACTGGCGTTGGATACACCCATCGAGTGGTA 349
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGATGT 3483
Db 350 GATCTTACCATACAAAATTTGCTCTCTCCCAAGACTCCGGGAGTTGATGATGT 401

RESULT 13
LOCUS BG418644 864 bp mRNA linear EST 23-OCT-2001
DEFINITION HVSMEK0023K16f Hordeum vulgare testae pericarp EST library
HVGDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0023K16f, mRNA
sequence.
ACCESSION BG418644
VERSION BG418644.1 GI:13324195
KEYWORDS EST.

```

SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 864)
AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Kannangara, G., von Wettstein, D., Akhunov, E., Chin, A., Choi, D.W., Fenton, R.D., Kianian, P., Otto, C., Simons, K., Zhang, D., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Oates, R. and Main, D.
TITLE Development of a genetically and physically anchored EST resource for barley genomics: Morex testa/pericarp cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases=535
Seq primer: AATTACCCCTCACTAAGGG
High quality sequence stop: 609.
Location/Qualifiers
1..864
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEK0023K16f"
/clone_lib="Hordeum vulgare testa/pericarp EST library
HVCNDA0013 (normal)"
/tissue types="testa/pericarp"
/lab host="TJC121"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
Plants were raised from seeds in a Controlled Environments growth chamber maintained in continuous light at 18°C, and testa and pericarp were dissected from developing kernels at Washington State University, Pullman, WA (Kannangara, von Wettstein). Total RNA was prepared, poly(A) RNA was purified, one cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids in the TJ Close lab at the University of California, Riverside (Akhunov, Chin, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 220 a 175 c 178 g 291 t
ORIGIN
Query Match 3.3%; Score 145.2; DB 12; Length 864;
Best Local Similarity 89.7%; Pred. No. 7.9e-20;
Matches 156; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 3310 GCAGCTTCATGAATTTGGCCAAACAAGGTACTAAGATTGACAGGTTCTTAATTTTCG 3369
Db 1 GCAGCTTCATGAATGCTGGCAAGCAAGGTGCTAAGTTGACAGGTTCTGAATTTTCG 60
Qy 3370 AATTGATGATCAATACTTGGGAAGCAAGTAATACCGGTTCGTTGGATCCACCCTCAAGTGG 3429
Db 61 AATTGATGATCAATATTGGGAAGCAAGTAATACTGGCCGTTTGATACATCCATCGAGTGG 120

Qy 3430 TAGATCTTATCATACAAAATTTGCTCTCTCTAAGACTCCTGGACTTGATGATGT 3483
Db 121 TAGATCTTACCATACAAAATTTGCTCTCTCTAAGACTCCGGAGTTGATGATGT 174

RESULT 14
BE495096/c
LOCUS BE495096
DEFINITION WHE1274_H05_P10Z5 Secale cereale anther cDNA library Secale cereale cDNA clone WHE1274_H05_P10, mRNA sequence.
ACCESSION BE495096
VERSION BE495096.1 GI:9661689
KEYWORDS EST.
SOURCE rye.
ORGANISM Secale cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Secale.
REFERENCE 1 (bases 1 to 593)
AUTHORS Anderson, O.D., Butler, E., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and Tong, J.C.
TITLE The structure and function of the expressed portion of the wheat genomes - Anther cDNA library from rye
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: StrataGene SK primer.
Location/Qualifiers
1..593
/organism="Secale cereale"
/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1274_H05_P10"
/clone_lib="Secale Cereale anther cDNA library"
/tissue type="Anther"
/dev stage="Adult plant before anthesis"
/lab host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Anthers were harvested and pooled from early meiosis to late meiosis. The tissue, total RNA, and poly(A) RNA were prepared (Butler, Ross and Gustafson) at University of Missouri, Columbia. A cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 165 a 125 c 114 g 189 t
ORIGIN

Query Match 3.3%; Score 144.8; DB 10; Length 593;
Best Local Similarity 90.1%; Pred. No. 9.6e-20;
Matches 155; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATTTGGCCAAACAAGGTACTAAGATTGACAGGTTCTTAATTTTCGAA 3371
Db 579 AGCTCGATGAATGCTGGCAAGCAAGGTGCTAAGTTTCACAAGGTTCTGAATTTTCGAA 520
Qy 3372 TTGATGATGAATACTTGGGAAGCAAGTAATACCGGTTCGTTGGATCCACCCTCAAGTGGTA 3431
Db 519 TTGATGATGAATATTGGGAAGCAAGGATCACTGGCCGTTGGATACACCCTCAAGTGGTA 460
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTCTAAGACTCCTGGACTTGATGATGT 3483

Search completed: March 15, 2003, 16:52:38
Job time : 5386.47 secs

```
Db 459 GATCTTACCATACAAAATTGCTCTCCTTAAGACTCCGGAGTTGATGATGT 408

RESULT 15
BE403629
LOCUS
DEFINITION BE403629 639 bp mRNA linear EST 21-JUL-2000
Triticum aestivum cDNA clone WHE0434_H01_002, mRNA sequence.
ACCESSION BE403629
VERSION BE403629.1 GI:9363097
KEYWORDS EST.
SOURCE Triticum aestivum
ORGANISM bread wheat.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 639)
Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
, P.S., Hsia, C.C., Kang, X., Lazo, G.R., Miller, R., Rausch, C.J.,
Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat
genomes
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: anders@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Strategene SK primer.
FEATURES
Location/Qualifiers
1..639
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0434_H01_002"
/clone_lib="Wheat etiolated seedling root cDNA library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Roots were
harvested. The tissue, total RNA, and poly(A) RNA were
prepared. A cDNA library was made, and the cDNA clones
were in vivo excised to give phagescript phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
```

BASE COUNT 167 a 148 c 170 g 154 t
ORIGIN

```
Query Match 3.3%; Score 144.8; DB 10; Length 639;
Best Local Similarity 90.1%; Pred. No. 9, 6e-20;
Matches 155; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAAATGTTGGCCAAACAGGCTACTTAAGATTGACAAGGTTCTTAATTTTGCAA 3371
|||||
Db 441 AGCTCGATGAAATGCTCGCAANGCAAGGTCCTAAGGTTGGCAAGGTTCTTAATTTTGCAA 500
|||||

QY 3372 TTGATGATGCAATACTGGAAGACGAATTACCGTCTGTTGGATCCACCCATCAAGTGGTA 3431
|||||
Db 501 TTGATGATGCAATTTGGAGACGATTAATGCGCGTTGGATACACCCATCGAGTGGTA 560
|||||

QY 3432 GATCTTATCATACAAAATTTGCTCTCTCTTAAGACTCCTCGACTTGATGATGT 3483
|||||
Db 561 GATCTTACCATACAAAATTTGCTCTCTCTCCCAAGACTCCGGAGTTGATGATGT 612
|||||
```


GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:16:40 ; Search time 10726.7 Seconds
(without alignments)
11777.621 Million cell updates/sec

Title: US-09-802-937-1
Perfect score: 4341
Sequence: 1 ctgcaggagattaattagg.....gtatatcttcattgtttt 4341

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- GENEID: 1
- 1: gb_ba.*
 - 2: gb_htg.*
 - 3: gb_in.*
 - 4: gb_om.*
 - 5: gb_ov.*
 - 6: gb_pat.*
 - 7: gb_ph.*
 - 8: gb_pl.*
 - 9: gb_pr.*
 - 10: gb_ro.*
 - 11: gb_sts.*
 - 12: gb_sy.*
 - 13: gb_un.*
 - 14: gb_vi.*
 - 15: em_ba.*
 - 16: em_fun.*
 - 17: em_hum.*
 - 18: em_in.*
 - 19: em_mu.*
 - 20: em_om.*
 - 21: em_or.*
 - 22: em_ov.*
 - 23: em_pat.*
 - 24: em_ph.*
 - 25: em_pl.*
 - 26: em_ro.*
 - 27: em_sts.*
 - 28: em_un.*
 - 29: em_vi.*
 - 30: em_htg_hum.*
 - 31: em_htg_inv.*
 - 32: em_htg_other.*
 - 33: em_htg_mus.*
 - 34: em_htg_pin.*
 - 35: em_htg_rod.*
 - 36: em_htg_mam.*
 - 37: em_htg_vrt.*
 - 38: em_sy.*
 - 39: em_htgo_hum.*
 - 40: em_htgo_mus.*
 - 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|--------------------|
| 2 | 3976.8 | 91.6 | 98782 | 2 | CNS08C85 | AL731756 Oryza sat |
| C 3 | 277.8 | 6.4 | 58527 | 8 | AB007649 | AB007649 Arabidops |
| C 4 | 229 | 5.3 | 1154 | 8 | RICADXA | DI0334 Oryza sativ |
| C 5 | 171.2 | 3.9 | 80115 | 8 | AB012248 | AB012248 Arabidops |
| C 6 | 164.4 | 3.8 | 387 | 8 | RICKN57 | DI0429 Rice mrna f |
| C 7 | 118.8 | 2.7 | 1085 | 8 | RICADKB | DI0335 Oryza sativ |
| C 8 | 115.6 | 2.7 | 13748 | 8 | AP003297 | AP003297 Oryza sat |
| C 9 | 112.8 | 2.6 | 772 | 8 | AY079340 | AY079340 Arabidops |
| C 10 | 112.8 | 2.6 | 904 | 8 | AY085188 | AY085188 Arabidops |
| C 11 | 112.8 | 2.6 | 979 | 8 | AY039945 | AY039945 Arabidops |
| C 12 | 109.2 | 2.5 | 124815 | 2 | AP005594 | AP005594 Oryza sat |
| C 13 | 109.2 | 2.5 | 138025 | 8 | AP003261 | AP003261 Oryza sat |
| C 14 | 109.2 | 2.5 | 169863 | 2 | CNS07YFX | AL713952 Oryza sat |
| C 15 | 108.8 | 2.5 | 126859 | 8 | OSJN00015 | AL606455 Oryza sat |
| C 16 | 108.8 | 2.5 | 155582 | 2 | OSJN00094 | AL606649 Oryza sat |
| C 17 | 108.4 | 2.5 | 100419 | 2 | OSIG00047 | AL732355 Oryza sat |
| C 18 | 107.6 | 2.5 | 100635 | 8 | AP003446 | AP003446 Oryza sat |
| C 19 | 107.6 | 2.5 | 144074 | 8 | AP003302 | AP003302 Oryza sat |
| C 20 | 106.4 | 2.5 | 156054 | 8 | AB023482 | AB023482 Oryza sat |
| C 21 | 104.8 | 2.4 | 129838 | 8 | AC079038 | AC079038 Oryza sat |
| C 22 | 103.6 | 2.4 | 170025 | 2 | OSJN00031 | AL606590 Oryza sat |
| C 23 | 102.6 | 2.4 | 1023 | 8 | AY086140 | AY086140 Arabidops |
| C 24 | 102 | 2.3 | 131983 | 8 | AC023240 | AC023240 Oryza sat |
| C 25 | 101 | 2.3 | 143959 | 8 | AC090055 | AC090055 Oryza sat |
| C 26 | 100 | 2.3 | 6735 | 2 | RIC3H3M | L28995 Oryza sativ |
| C 27 | 100 | 2.3 | 129845 | 2 | AP004165 | AP004165 Oryza sat |
| C 28 | 99.4 | 2.3 | 121615 | 2 | AP003981 | AP003981 Oryza sat |
| C 29 | 99.4 | 2.3 | 146670 | 8 | AP003250 | AP003250 Oryza sat |
| C 30 | 99 | 2.3 | 142596 | 8 | AP003314 | AP003314 Oryza sat |
| C 31 | 99 | 2.3 | 166753 | 8 | AP002866 | AP002866 Oryza sat |
| C 32 | 98.4 | 2.3 | 159893 | 8 | OSJN00095 | OSJN000171 |
| C 33 | 98.4 | 2.3 | 189834 | 8 | CNS08C8F | AL662974 Oryza sat |
| C 34 | 96.8 | 2.2 | 108805 | 2 | CNS08C89 | AL731875 Oryza sat |
| C 35 | 96.8 | 2.2 | 150238 | 2 | AP003453 | AL831796 Oryza sat |
| C 36 | 96.6 | 2.2 | 151100 | 8 | AP003453 | AP003453 Oryza sat |
| C 37 | 96.2 | 2.2 | 136553 | 2 | AC093953 | AC093953 Oryza sat |
| C 38 | 95.8 | 2.2 | 132407 | 8 | OSJN00117 | AP007002 Oryza sat |
| C 39 | 95.2 | 2.2 | 139566 | 2 | AP004161 | AP004161 Oryza sat |
| C 40 | 95.2 | 2.2 | 147954 | 2 | AP003983 | AP003983 Oryza sat |
| C 41 | 95.2 | 2.2 | 165394 | 8 | AC025296 | AC025296 Oryza sat |
| C 42 | 94.8 | 2.2 | 160174 | 2 | AP005103 | AP005103 Oryza sat |
| C 43 | 94.8 | 2.2 | 162010 | 2 | AP004121 | AP004121 Oryza sat |
| C 44 | 94.6 | 2.2 | 102119 | 2 | AP003919 | AP003919 Oryza sat |
| C 45 | 93.2 | 2.1 | 122559 | 2 | AP004301 | AP004301 Oryza sat |

ALIGNMENTS

RESULT 1
AB041773
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AB041773
Oryza sativa Adk-a gene for adenylyate kinase, partial cds.
AB041773.1 GI:7630192
adenylyate kinase.
Oryza sativa DNA.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 4341)
Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S.
Oryza sativa Ak-a gene,promoter region and partial ORF

JOURNAL Published Only in DataBase (2000)
REFERENCE 2 (bases 1 to 4341)
AUTHORS Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Hiromitsu Fukuzawa, Nissan Chemical
Industries,Ltd, Shiroaka Research Station of Biological Science,
1470 Shiroaka, Minamisaikama, Saitama 349-0294, Japan
(E-mail:fukuzawa@nissaichem.co.jp, Tel:+81-480-92-2513)
Location/Qualifiers

FEATURES

source

1. 4341
/organism="Oryza sativa"
/db_xref="taxon:4530"
join(1475..1577,2841..2986,3074..3193,3314..3481,
4126..4215)
/gene="Adk-a"
join(1475..1577,2841..2986,3074..3193,3314..3481,
4126..4215)
/gene="Adk-a"
/codon_start=1
/product="adenylate kinase"
/protein_id="BAA94761.1"
/db_xref="GI:7630193"
/translation="MAANLREDVPSMELMTTELLRMKSSKPKRVLVGPQGGKGTQ
SPLIKDFLCHLATQDMLEAAVAAKTPIGIKAKEAMDKELVSDLVVGIIDEAMKK
TSCOKGFLDFPRVTVOAKQLDEMLAKOGTKIDVLFNFIADAILERITGRWIHPS
SGRSYHTKFAPPKTPGUDVDTGELIQRKDDTAAVLKSLEAFHVQTKP"
BASE COUNT 1284 a 758 c 880 g 1419 t
ORIGIN

Query Match 100.0%; Score 4341; DB 8; Length 4341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGAGGAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGACCCCTGTGTGTT 60
DB 1 CTGAGGAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGACCCCTGTGTGTT 60
QY 61 AATCAATCGGGGTTCGTGGACATGCTTTTGCAGGAAAATTAAGCAAGAAAATTAAG 120
DB 61 AATCAATCGGGGTTCGTGGACATGCTTTTGCAGGAAAATTAAGCAAGAAAATTAAG 120
QY 121 AAGAATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCAATTTCAAGTGGTTC 180
DB 121 AAGAATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCAATTTCAAGTGGTTC 180
QY 181 TCTTGTAACCACTGTTAGAAATACATAGACAGTGCACCGTTGATGGCTCCTATTGG 240
DB 181 TCTTGTAACCACTGTTAGAAATACATAGACAGTGCACCGTTGATGGCTCCTATTGG 240
QY 241 CTCGTGTGATCTGACTTGTGTCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATTA 300
DB 241 CTCGTGTGATCTGACTTGTGTCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATTA 300
QY 301 CCGAAAAACCAAGATTATTTCTATTCACCTCAGGGTAAATTTGCTGAACTATGCAATG 360
DB 301 CCGAAAAACCAAGATTATTTCTATTCACCTCAGGGTAAATTTGCTGAACTATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTCTCAAAATGAAATTTGAGTCCAA 420
DB 361 AATACAAATTCGCAAAATATCATGTTATCTATCTCTCAAAATGAAATTTGAGTCCAA 420
QY 421 CTGAGACTGCAATAGATTCTTTTCAAAAAGAAATTAATATTTTCTTTTTCATGAAA 480
DB 421 CTGAGACTGCAATAGATTCTTTTCAAAAAGAAATTAATATTTTCTTTTTCATGAAA 480
QY 481 CGCAATTCACCGTTTCGAGAAATATGCTGCATAAATTAAGTGTAGTGCAGAAACAAA 540
DB 481 CGCAATTCACCGTTTCGAGAAATATGCTGCATAAATTAAGTGTAGTGCAGAAACAAA 540
QY 541 ATTAATATCACATAAAAAGAGTTGTTAATTAACAACCATGTTTCGTACTCAACTCT 600
DB 541 ATTAATATCACATAAAAAGAGTTGTTAATTAACAACCATGTTTCGTACTCAACTCT 600

QY 601 AATTGTAAAAATCTTATTTCAGTCACAAAATTCCAATTTCCAAATTAAGAAAAATAAAGCT 660
DB 601 AATTGTAAAAATCTTATTTCAGTCACAAAATTCCAATTTCCAAATTAAGAAAAATAAAGCT 660
QY 661 AGACGGCTAGAGCCACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAGCAAAAATA 720
DB 661 AGACGGCTAGAGCCACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAGCAAAAATA 720
QY 721 TGATGGTTTAAATATGATGTTTTTAAATAACTTTTACATAAAATTTCTTTAGGAAACA 780
DB 721 TGATGGTTTAAATATGATGTTTTTAAATAACTTTTACATAAAATTTCTTTAGGAAACA 780
QY 781 TATCATTTAATGTTTGAAGAAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGGAAA 840
DB 781 TATCATTTAATGTTTGAAGAAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGGAAA 840
QY 841 CAACAGAAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTCTAGTGGAGTTGATTTTCAAC 900
DB 841 CAACAGAAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTGGAGTTGATTTTCAAC 900
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTATAATTTATAAATCT 960
DB 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTATAATTTATAAATCT 960
QY 961 TGAATAAAATATTTATTTGAATTTTTTAAACAATGATGCATAAAATTTATTTTAAAAAAC 1020
DB 961 TGAATAAAATATTTATTTGAATTTTTTAAACAATGATGCATAAAATTTATTTTAAAAAAC 1020
QY 1021 ACACCAATTTAAACCTTTTAAAGCATCTTAATAGAAAACGAGGAAGTTTAAAGATTCACC 1080
DB 1021 ACACCAATTTAAACCTTTTAAAGCATCTTAATAGAAAACGAGGAAGTTTAAAGATTCACC 1080
QY 1081 GAAGTGTTCGATAATGAAAAATGGGTGGATTAGAAATTTGGTAAATGAATCAGGGTTAG 1140
DB 1081 GAAGTGTTCGATAATGAAAAATGGGTGGATTAGAAATTTGGTAAATGAATCAGGGTTAG 1140
QY 1141 GATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTCTTTTG 1200
DB 1141 GATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTCTTTTG 1200
QY 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCTTGCCTGGGCAGA 1260
DB 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCTTGCCTGGGCAGA 1260
QY 1261 AGCGTCTTTTTCGTGGAAAAAATCTGTTTAAAGGAAAACAGAAAGAGCCGAGCTTT 1320
DB 1261 AGCGTCTTTTTCGTGGAAAAAATCTGTTTAAAGGAAAACAGAAAGAGCCGAGCTTT 1320
QY 1321 GTTGTCAACCTCTCACAGAAACGAAACAAAAGCCCCACCACTTAAACCTCTCGATCC 1380
DB 1321 GTTGTCAACCTCTCACAGAAACGAAACAAAAGCCCCACCACTTAAACCTCTCGATCC 1380
QY 1381 GACCGAGACTCTCCATTTACGCGGCGCACGCGGAGAGCACGCGAGTTCGTTCGAC 1440
DB 1381 GACCGAGACTCTCCATTTACGCGGCGCACGCGGAGAGCACGCGAGTTCGTTCGAC 1440
QY 1441 GAACAAGCTAGTGCAGTAGTGTCTGCGGAGATGGCGGCGAACCTTGGAGAGCGTGC 1500
DB 1441 GAACAAGCTAGTGCAGTAGTGTCTGCGGAGATGGCGGCGAACCTTGGAGAGCGTGC 1500
QY 1501 GTCGATGGAGCTGATGACGAGTGTCTCGCGCGCATGAAGTGCAGTCCAAAGCCGACAA 1560
DB 1501 GTCGATGGAGCTGATGACGAGTGTCTCGCGCGCATGAAGTGCAGTCCAAAGCCGACAA 1560
QY 1561 GCGGCTCATCTCGTCGGTAAACGCCCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCTC 1620
DB 1561 GCGGCTCATCTCGTCGGTAAACGCCCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCTC 1620
QY 1621 TCGGTGTTTGGTTGTTGGTGGATCCGGTTCGCGGCGCGCTGCTGCCCTGGTCCG 1680
DB 1621 TCGGTGTTTGGTTGTTGGTGGATCCGGTTCGCGGCGCGCTGCTGCCCTGGTCCG 1680
QY 1681 GAGATGGGAAAAGATTGCTGTGGCTTGTGCGGATGTTGCGTTAGATCCGTGCAAAAGCTG 1740


```
Qy 3901 CTTGTTTAAAGTGGCCATATAATAGATACATGACATGTTTTTTTGTAGCAGAGATTTGTATA 3960
Db 3901 CTTGTTTAAAGTGGCCATATAATAGATACATGACATGTTTTTTTGTAGCAGAGATTTGTATA 3960
Qy 3961 TTGTTTCTATTGCTTCCAAATTAAGAGATATCTCTTTTAAAGCAATGATTTTCATGTTGGAC 4020
Db 3961 TTGTTTCTATTGCTTCCAAATTAAGAGATATCTCTTTTAAAGCAATGATTTTCATGTTGGAC 4020
Qy 4021 ATATTGTCCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTTAAATTTGTTAAT 4080
Db 4021 ATATTGTCCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTTAAATTTGTTAAT 4080
Qy 4081 ATATTGTCCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTTAAATTTGTTAAT 4140
Db 4081 ATATTGTCCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTTAAATTTGTTAAT 4140
Qy 4141 TTAATTCAGGAAGATGACACAGCTGCAGTATTAAGTCAAGGCTTGAAGCTTCCAC 4200
Db 4141 TTAATTCAGGAAGATGACACAGCTGCAGTATTAAGTCAAGGCTTGAAGCTTCCAC 4200
Qy 4201 GTACAACTAAGCCCTGATGTTTCTTTTAGCACTACGTTTTTTTAAATATTCAGATATCT 4260
Db 4201 GTACAACTAAGCCCTGATGTTTCTTTTAGCACTACGTTTTTTTAAATATTCAGATATCT 4260
Qy 4261 TTTAGGATGATGCTACTTCAGTTAAGCGGATTCCTTCAGTTGCAATACAGTGTTC 4320
Db 4261 TTTAGGATGATGCTACTTCAGTTAAGCGGATTCCTTCAGTTGCAATACAGTGTTC 4320
Qy 4321 TGTATATCTTTCATTTGTTTTT 4341
Db 4321 TGTATATCTTTCATTTGTTTTT 4341

RESULT 2
CNS08C85 98782 bp DNA linear HTG 10-JUL-2002
LOCUS Oryza sativa chromosome 12 clone OJ1220_D01, *** SEQUENCING IN
DEFINITION PROGRESS **, in ordered pieces.
ACCESSION AL731756
VERSION AL731756.3 GI:21732197
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE
AUTHORS Choise, N., Orjeda, G., Cattolico, L., Demange, N., Wincker, P., Seguren, B., Pelletier, E., Scarpelli, C., Salanoubat, M., Weissenbach, J., and Quetier, F.
TITLE Oryza sativa chromosome 12 sequencing
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 98782)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]
COMMENT On Jul 11, 2002 this sequence version replaced gi:20372820.
IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence.
Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
```

```
source 1..98782
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OJ1220_D01"
/clone_lib="Monsanto"
BASE COUNT 30065 a 19192 c 20278 g 29238 t 9 others
ORIGIN
Query Match 91.6%; Score 3976.8; DB 2; Length 98782;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 4254; Conservative 0; Mismatches 72; Indels 44; Gaps 19;
Qy 1 CTGCAGGAAGATTAATTTAGGTGGACACACAAACCTGTTGGTGGACGCCCTGTTGTT 60
Db 63789 CTGCAGGAAGATTAATTTAGGTGGACACACAAACCTGTTGGTGGACGCCCTGTTGTT 63848
Qy 61 AATCAACTGGGTGTTGCTGTTGGACATGTTTTCAGGAAAAATTAAGCAAGAAAAATTAAG 120
Db 63849 AATCAACTGGGTGTTGCTGTTGGACATGTTTTCAGGAAAAATTAAGCAAGAAAAATTAAG 63908
Qy 121 AAGAACTCTCAAGCTGACATGAGAAAAACGTAATCCAAATGGAAGCGAAATTTCAAGTCGTT 180
Db 63909 AAGAACTCTCAAGCTGACATGAGAAAAACGTAATCCAAATGGAAGCGAAATTTCAAGTCGTT 63968
Qy 181 TCTTGTAATACCAATGTTTGAATATACATAAGACAGTGCACAC - GTTTGATGGCTCTTATTG 239
Db 63969 TCTTGTAATACCAATGTTTGAATATACATAAGACAGTGCACAC - GTTTGATGGCTCTTATTG 64028
Qy 240 GCTCGTGTGATGACTGTTGTGCACAAAGCATCAAAATGCTTCTTTGGAGTATCTTTATT 299
Db 64029 GCTCGTGTGATGACTGTTGTGCACAAAGCATCAAAATGCTTCTTTGGAGTATCTTTATT 64088
Qy 300 ACCGAAACCCCAAGATTAATTTCTATTCACCTCAGGGTAAATTTGCTGAACTATGCAAT 359
Db 64089 ACCGAAACCCCAAGATTAATTTCTATTCACCTCAGGGTAAATTTGCTGAACTATGCAAT 64148
Qy 360 GAATACAAATTCGCAAAATATCATGTTTATCTATCTTGTCAAATTTGAAATTTGAGTCCA 419
Db 64149 GAATACAAATTCGCAAAATATCATGTTTATCTATCTTGTCAAATTTGAAATTTGAGTCCA 64208
Qy 420 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATAATTTTTCATGAA 479
Db 64209 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATAATTTTTCATGAA 64268
Qy 480 ACCCAATTCACCGTTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAA 539
Db 64269 ACCCAATTCACCGTTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAA 64328
Qy 540 AATTAATATCACATAAAAAAGAGGTTGTTTAATTAATAAACCATTGTTTCGTACTACAATC 599
Db 64329 AATTAATATCACATAAAAAAGAGGTTGTTTAATTAATAAACCATTGTTTCGTACTACAATC 64388
Qy 600 TAATTTGTAATTTCTTATTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAAATTAACG 659
Db 64389 TAATTTGTAATTTCTTATTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAAATTAACG 64448
Qy 660 TAGACGGCTAAGCCCAACCCATCTTAAGGCTTAAGTTTCGAGAGGTGAAGTACGCACGAAAAAT 719
Db 64449 TAGACGGCTAAGCCCAACCCATCTTAAGGCTTAAGTTTCGAGAGGTGAAGTACGCACGAAAAAT 64508
Qy 720 ATGATGGTTTAAATATGATTTTTTTTAAATTAACCTTTTACATAAATTTCTTTAGGAAC 779
Db 64509 ATGATGGTTTAAATATGATTTTTTTTAAATTAACCTTTTACATAAATTTCTTTAGGAAC 64568
Qy 780 ATATCATTTTAAATGTTTGAAGAAACCTGCAATAGAAAACTAAGAACGATGAGTTGGAA 839
Db 64569 ATATCATTTTAAATGTTTGAAGAAACCTGCAATAGAAAACTAAGAACGATGAGTTGGAA 64626
Qy 840 ACAGAGAAAAACACAGCGCTTAAGGCTTCTGATCTCTGATGAGGTTGATTTTCAA 899
```

| | | | |
|----|-------|---|-------|
| Db | 64627 | ACAAGAGAAAAACACAGCCTTAAGGCTTCTTGATGCTCTAGTTGGAGGTGTGATTTTCAAA | 64686 |
| Qy | 900 | CGCATGATAAACGAGAAAGCTCATTAGCACATATTACTTTAGATATATTATAATTATAAAC | 959 |
| Db | 64687 | CGCATGATAAACGAGAAAGCTCATTAGCACATATTACTTTAGATATATTATAATTATAAAC | 64746 |
| Qy | 960 | TTGAAAAAATATTATTGTAATTTTAAACAATGATCGATAAATATTATTTTAAAAAA | 1019 |
| Db | 64747 | TTGAAAAAATATTATTGTAATTTTAAACAATGATCGATAAATATTATTTTAAAAAA | 64806 |
| Qy | 1020 | CACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGAAACGAGGAAGTTAAAGATTCCAC | 1079 |
| Db | 64807 | CACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGAAACGAGGAAGTTAAAGATTCCAC | 64866 |
| Qy | 1080 | CGAA--GTGTTTGGATAAT-GAAAAATGGGTCGGGATTAGAAATCGTAATGAATCAGG | 1135 |
| Db | 64867 | CGAAGGTGTGTTGGATAATGGAANAATGGGTCGGGATTAGAAATCGTAATGAATCAGG | 64926 |
| Qy | 1136 | GTTAGGATTAATAATTAATAATGAAGAGGAGAATGAATGGTTAGAGTTTAAATCTGTCT | 1195 |
| Db | 64927 | GTTAGGATTAATAATTAATAATGAAGAGGAGAATGAATGGTTAGAGTTTAAATCTGTCT | 64986 |
| Qy | 1196 | TTTTGGTGGGTAGAAAAATTATTGTCATACACTCCCGAGAGCGGTGCTGCTCGGTGG | 1255 |
| Db | 64987 | TTTTGGTGGGTAGAAAAATTATTGTCATACACTCCCGAGAGCGGTGCTGCTCGGTGG | 65046 |
| Qy | 1256 | GCAGAAAGCGTCTTTTTCGTTGGAATAAAAAAACTGCTTTAAAAAGGAAACAGAAAGAGCCCA | 1315 |
| Db | 65047 | GCAGAAAGCGTCTTTTTCGTTGGAATAAAAAAACTGCTTTAAAAAGGAAACAGAAAGAGCCCA | 65106 |
| Qy | 1316 | GCCTTT-GTTGTCACCGTCTCACAGAAACGAAACAAAAGCCCCACACCTAAACCTCCT | 1374 |
| Db | 65107 | GCCTTGGTGTCTACCGTCTCACAGAAACGAAACAAAAGCCCCACACCTAAACCTCCT | 65166 |
| Qy | 1375 | CGATCCGACGAGACTCTCCATTTTCAGCGCGGCACGCGAGAGACACGGCAGCGAGTTTC | 1434 |
| Db | 65167 | CGATCCGACGAGACTCTCCATTTTCAGCGCGGCACGCGAGAGACACGGCAGCGAGTTTC | 65226 |
| Qy | 1435 | GTCCAGCAACAAAGCTTAGTCAGTAGTTGTTGTCGCGGAATGCGCGCGAACTCGAGAGA | 1494 |
| Db | 65227 | GTCCAGCAACAAAGCTTAGTCAGTAGTTGTTGTCGCGGAATGCGCGCGAACTCGAGAGA | 65286 |
| Qy | 1495 | CGTCCCGTCGATGGAGCTGATGACGGAGCTGCTCCGCGGCATGAAGTGAGCTCCAAGCC | 1554 |
| Db | 65287 | CGTCCCGTCGATGGAGCTGATGACGGAGCTGCTCCGCGGCATGAAGTGAGCTCCAAGCC | 65346 |
| Qy | 1555 | CGACAAGCGCGTCATCCTCGTCGCTTAACGCCCGCTCTTTCTTCTTCTCTCTCTCTC | 1614 |
| Db | 65347 | CGACAAGCGCGTCATCCTCGTCGCTTAACGCCCGCTCTTTCTTCTTCTCTCTCTCTC | 65406 |
| Qy | 1615 | TCCTCTCCGTCGGTTTGTTTCGTTGGTGTGGATCCGGTTCCGCGCGCGTGCCTCCCT | 1674 |
| Db | 65407 | TCCTCTCCGTCGGTTTGTTTCGTTGGTGTGGATCCGGTTCCGCGCGCGTGCCTCCCT | 65466 |
| Qy | 1675 | GGTCCGGAGATGGGAAAGATTGCTGTGGCTTGTCTGGGGAATGTTGGGTTAGATCCGTGCA | 1734 |
| Db | 65467 | GGTCCGGAGATGGGAAAGATTGCTGTGGCTTGTCTGGGGAATGTTGGGTTAGATCCGTGCA | 65526 |
| Qy | 1735 | AAGCTGTCCCTTCTCTTTTGGGTCAAACTCTGGGTTCTCTTCGCTCTCTCTCGT | 1794 |
| Db | 65527 | AAGCTGTCTCCCTTCTCTTTTGGGTCAAACTCTGGGTTCTCTTCGCTCTCTCTCGT | 65586 |
| Qy | 1795 | GTTTTTGC AAAACAAAATGTTTGGCTTCGAGGGAATCTTGTAAATTTACTTCTGTT | 1854 |
| Db | 65587 | GTTTTTGC AAAACAAAATGTTTGGCTTCGAGGGAATCTTGTAAATTTACTTCTGTT | 65646 |
| Qy | 1855 | TCACACTTTGGAAACTCGCGCAATTTTGTGATGTCGTAATTTGGTGTGTTCTTTCGAACTT | 1914 |
| Db | 65647 | TCACACTTTGGAAACTCGCGCAATTTTGTGATGTCGTAATTTGGTGTGTTCTTTCGAACTT | 65705 |
| Qy | 1915 | GCGAATTGAGTCGATTGGGTTTTTCCATTTGACGGAAAGGATATTAGTGACAG-AAATCAAG | 1973 |
| Db | 65706 | GCGAATTGAGTCGATTGGGTTTTTCCATTTGACGGAAAGGATATTAGTGACAGAAATCAAG | 65765 |

| | | | |
|----|-------|---|-------|
| Qy | 1974 | CGAAAGAAATATTTATATGATCGGTGTTTCGACCCATAATAGGAATTTTCAGCGGTTTAAG | 2033 |
| Db | 65766 | CGAAAGAAATATTTATATGATCGGTGTTTCGACCCATAATAGGAATTTTCAGCGGTTTAAG | 65825 |
| Qy | 2034 | GTATAGAAATTTATTCGTATTCAGTAGTAGTTTGAAATTTGGAGGTTGCAATTTTTTGTTTC | 2093 |
| Db | 65826 | GTATAGAAATTTATTCGTATTCAGTAGTAGTTTGAAATTTGGAGGTTGCAATTTTTTGTTTC | 65885 |
| Qy | 2094 | TCACACATGTTATTTTAAATTTGTCATCGGCCAATGCTGTTTGAGGATGAGCAAGAAAAACA | 2153 |
| Db | 65886 | TCACACATGTTATTTTAAATTTGTCATCGGCCA--TGTGTTGAGGATGAGC-AGAAAAACA | 65942 |
| Qy | 2154 | ACAGGCGGTTTCGTGTTTTTTTGTCTGATGTGGAAAACTTTTAAAAATAACCAAAACAGAA | 2213 |
| Db | 65943 | ACAGGCGGTTTCGTGTTTTTTTGTCTGATGTGGAAAACTTTTAAAAATAACCAAAACAGAA | 66002 |
| Qy | 2214 | AGGTGCAACTTTATATGATTCGATAAAGGTTTGAAAAATGTCACTCGGAATAAAAAACAAG | 2273 |
| Db | 66003 | AGGTGCAACTTTATATGATTCGATAAAGGTTTGAAAAATGTCACTCGGAATAAAAAACAAG | 66062 |
| Qy | 2274 | AGATAGACGTTATPACCTCTTGTGTTTCAATATTCACACTACTAAACCAATCGATCTGTATGG | 2333 |
| Db | 66063 | AGATAGACGTTATPACCTCTTGTGTTTCAATATTCACACTAGTAAACCAATCGATCTGTATGG | 66122 |
| Qy | 2334 | TTATGTAACCTGTGGACAAACAAATTTCAAAATAAACTTGTTCTGTGGCAATGCTCT | 2393 |
| Db | 66123 | TTATGTAACCTGTGGACAAACAAATTTCAAAATAAACTTGTTCTGTGGCAATGCTCT | 66182 |
| Qy | 2394 | GCCTCAGCAATGCACCTACTCTAAATTTACTGGATCATCAAAACAAATCATATGGAAGGTT | 2453 |
| Db | 66183 | GCCTCAGCAATGCACCTACTCTAA--TTTACTGGTCATCAAAACAAATCATATGGAAGGTT | 66240 |
| Qy | 2454 | TATTCATATCGTCTTTTACGTTCTATGAAGTATAGCTAAATTTTCAAACTCTCTGCACCTAC | 2513 |
| Db | 66241 | TATTCATAT--GTTCTTTTACCTTCAAGAAGTATAGCTAAATTTTCAAACTCTCTGCACCTAC | 66298 |
| Qy | 2514 | CTAATTTTATGTGGACCGAATCTAGTCATCGTCTCGGTTAGCTATCTCCAAATTTGATGCG | 2573 |
| Db | 66299 | CTAATTTAAGTGGACCGAATCTAGTCATCGTCTCGTCTCTTAGCTATCTCCAAATTTGATGCG | 66358 |
| Qy | 2574 | GTGTTCTTAATTCATGATGCTGTCACTGAGA-ACGTGGTGTGTTTAAACAGCAATTTGAAGAAT | 2632 |
| Db | 66359 | TTGTTCTTAATTCATGATGCTGTCACTGAGACACATCGTGTGTTGTTTAAACAGCAATTTGAAGAAT | 66418 |
| Qy | 2633 | GCCAG-TACACCCATAATGTTATTGAGGATAGATAGTG-----A2672 | |
| Db | 66419 | GCCAGTTTACACCCATAATGTTATTGAGGATAGATAGTTTGCAAACTTGATATCCATA66478 | |
| Qy | 2673 | CAGTGTAAACATAGACAGCAAGTGTATGCAATTCATGCTTTAGATAGTTTATAGCATCTGGCA2732 | |
| Db | 66479 | GAGCAAAAGTGATGCAAAATTTTGATGCTTTTATGGATAGTTTAAATAGGCATCTGGCAG66538 | |
| Qy | 2733 | GAGCTTTGAATATTGATG-TCTAGAATGGTAAGACAATCTCGGTGTTCTTAAAAAATAA2791 | |
| Db | 66539 | AAAGCTTGAATATTGATGTTCTAGAAUGGTAAGACAATCTCTCGGTGTTCTTAAAAAATAA6598 | |
| Qy | 2792 | AAAGGTAAGACATACAATTTTGGACATCCCTTTATTTTACTAAATTTTAGGTCCA66581 | |
| Db | 66599 | AAAGGTAAGACATACAATTTTGGACATCCCTTTATTTTACTAAATTTTAGGTCCA6658 | |
| Qy | 2852 | TGCGAAAGGGAACACAGTCACCGCTGATTAAAGGATGAAATTTTGTGTTGTCATTTTAGCC2911 | |
| Db | 66659 | TGCGAAAGGGAACACAGTCACCGCTGATTAAAGGATGAAATTTTGTGTTGTCATTTTAGCC66718 | |
| Qy | 2912 | ACTGGTGATATGTTGAGGCTGCATGAGCTGCTAAACTCCACTCGGATTAAGGCTAAA2971 | |
| Db | 66719 | ACTGGTGATATGTTGAGGCTGCATGAGCTGCTAAACTCCACTCGGATTAAGGCTAAA66778 | |
| Qy | 2972 | GAAGCTATGGACAAGGTAGTTTTTAAAGAAACATATAGCAACAGAAATTAATAACGACAGG3031 | |
| Db | 66779 | GAAGCTATGGACAAGGTAGTTTTTAAAGAAACATATAGCAACAGAAATTAATAACGACAGG66838 | |


```
/product="4-coumarate-CoA ligase-like protein"
/notes="CDS is reported in Acc# AB023035
gene id:K9H21.8"
/number=2
/evidence=not experimental
complement(625..1680)
/product="4-coumarate-CoA ligase-like protein"
/notes="CDS is reported in Acc# AB023035
gene id:K9H21.8"
/number=1
/evidence=not experimental
complement(join(2166..2408,2476..2734,2838..3121,
3189..3278,3360..3528,3611..4245))
/notes="gene id:MLE2.2"
/codon_start=1
/evidence=not experimental
/product="auxin-independent growth promoter-like protein"
/protein_id="BAB08804.1"
/db_xref="GI:9758280"
/translation="WAKSMRTIKNPFFTSHPSPFHLILLFFPLKSSPRVNLHHH
QRCPPQFLLISLVFSGISFLTSLDTGSTCVSTSSASLKFIFISDVDRHQI
LTSLVAGASTLUPPARGGHSGNMTEBEKFPKOPGEGYKCLDFSLSYKKSASV
SEKRFLLVVGGLNQORNIQVAVIAMLEALVVPVQVNRVMGDESEFSLF
DVEHKTLTSDVRIVSSLPSTHLMRSQTIENQIPWDVSPVMIRAKYFKQINLEGHLV
EKGLDSKLAKNLPDQLKRCVAFHARFAPIENLGNKLTIRRMTEGYVIAHLRL
KGVNVRTGLGSEFDRIIAETFTSQPYLTGRNLTYTERRLAGFCPLNVEIYA
RLKALGPSNASIYAGEPEGSGRALEPLAKFSNLVTKETLAHKGELLPTNRS
ALAADIYVLSDDVFIPISHGGMKAKQGNRAVYGRKFMIPNKRAMLPMENSYS
DAELSLTKLHKRKGQGPESRRGRDRDVIAYVPECMCHRRKRSVGGF"
complement(join(4928..5026,5118..5207,5299..5466,
5572..5691,5787..5932,6354..6471))
/notes="gene id:MLE2.3"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB08805.1"
/db_xref="GI:9758281"
/translation="WATGGAADLEDVQTVLMSBELRLKCSQKPKRILFIFGPGS
GKGTSPVVKDYCLHSTGDMRAAVASKTPGLVKAKEAMEKGLVSDLVVGIID
EAMNPKCKOGFILDGFPRTVTAQELDEMILKRGTEIDKVLNFAIDDAILERITGR
WIHPSSGRSYHTKFAPPKTPGVDDITGBPLQRKDDNADVLKSLAAAFHQTPQVIDY
YAKKAVLTNIOAEKAPQEVTSVKALS"
complement(join(6918..7036,7117..7245,7475..7648,
7750..7993,8083..8219,8312..9422))
/notes="contains similarity to receptor-like protein
kinase
gene id:MLE2.4"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB08806.1"
/db_xref="GI:9758282"
/translation="WGLFKFIFLLSLLMSFYSLSSQLQASQAQVLLQLKKHLEYPQO
LESWDYHRTNFCYLOATSMNITCFNSVSESLNIPGDKSEKAKSPFGFAISNVLSD
GTSIESFVTLTSLRLKSLRVLTIASLGIWGRLEPKELHLSLEYLDLNNFLFGSVPPK
LSTMVLETFRFDHNFNGTLPSWFDYVGLKSLVLSFKSNKLSGELHSLLSLSTIEYI
DLRANSLGSLPDDLCKCSKLAFIDISDKLTGKLPRCLSSKQDIALRFNGNCILSLEK
QHPSPFCYKVRAAKAEAKAEAEANESGKRWKKGALLIGLIVGSMVLVLCVY
FILLRKGTGKHVHNTVQDDNHTTIGFSSEILSNARTISETSIFGSEDLFVCKQSL
BEIVKATNFDMTILGESSLYGTLYKGNLENGTKVAIRCLPSSKKYSIRNLKLRDL
LAKLRHNLVCLLGHICDGGKDDYSVEKFLIYEIYIPNGFQCLSDNSGSGKMMWS
ERNVLTVGAKAVHFLHTGVIYPGFSSNRLKTNVLLNQHRFAKLSDYGNRKIMANVKA
RRLCLQMTDSTINLHTWITICIRKRSISITRRTGIVGERRRKEGSDSTSHLSKRIT
DKSDNSDEOMCT"
join(12169..13370,12719..12950,13056..13179,13364..13420,
13649..13840,13919..13996,14120..14302,14505..14585,
14791..14856,14952..15044,15143..15268,15478..15630)
/notes="gene id:MLE2.5
similar to unknown protein
sp|P54122"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08807.1"
/db_xref="GI:9758283"

/translation="MMKPASLQGFSSHASSIYSDVRRPATTFFSKMAAFSALSCLPYT
FTFQSSRIKSTVSCVTSASPSTSSSTPRRSRLEGVSKMEDSVKRWEOFY
EGTDPPPLRIILIGLGEIIGMNCMLVGNIDRYILIDAGIMFPDYDEPGIOLIMPTGTF
IRRWKKEIAVITHGHEHIGALPWIIPALDNTPIFASFTWELIKKLKEHIGIFV
OSRLKTESTRERFMAEPLEPIVTHSIIDCSGLFLRCADGNILHVNMAKIGVITTO
KGVDFREALEELSKEGVTLNMSDSTNVLSFGRTISEKVDKAPILNMAKIGVITTO
FASNIHRLGTSKAAADITGRKLVFGMSLRTYLEAAWRDGRAPIDPSLSIKVSDIEAY
APKDLLVITTSQAEPRAAANLASYSGHAFKLTKEIDILYSKAVIPGNSRVMKMN
RIADIGNIIMGKNEMLTSHAYRGELVLKIVKQHFILPHGELFLFKHLELLGKST
GIRHTVLSLSPKTYELMVDK"
join(16291..16714,16783..17258,17369..17461)
/notes="gene id:MLE2.6
pir||T01257"
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08808.1"
/db_xref="GI:9758284"
/translation="MEIMRPGVSENTLKGRITTRCMWLDKGRLLDALHKAHAALS
SCPVTCPLSHMERTVSEVLKIVKYSKSPVIAIATENPMARVADVEARLSGDPS
VSGVVAALRKVVEGNDKRSRAKAPSQEASPKVDRTLEDDIIDSARLLAEETAST
YTEEDVTPVSGSSSEDDFWKSFINPSSPSESTENNNKVAOTPEKAREGSRUDD
ELADASETSKSPKRVKKNKPEEIKKIVMRGELHRSFQVVVKRMAWBEISSNL
SAEGINRSPGCKSLWASLIQYKESKADERSKTSWPHFEDMNNILSELGTPAS"
join(18617..18789,18882..19035,19112..19208,19304..19392,
19521..19589,19790..19849)
/notes="gene id:MLE2.7
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08809.1"
/db_xref="GI:9758285"
/translation="MQSGILVCLYTNLDPLOQLQKMPKRTDRSNVLDKKTHLA
RLNVSEGGKVLKRGEGMRQFMCNIGCELFVCYRAENLETASFIIYVDGALSVA
AAETNPQDAPVPCISQLDGGLVOVAJEVEDRAORSATRVNADDDVTVTAAPARGE
ANNELEFNGVRLGRLSQMTLQRGNSKSLVVEDLSARQYVEKLEAVP"
complement(20641..22173)
/notes="gene id:MLE2.8"
/codon_start=1
/evidence=not experimental
/product="cytochrome P450-like protein"
/protein_id="BAB08810.1"
/db_xref="GI:9758286"
/translation="WEMLNAILLIFPIGVLFIPPTTLKAKTASPSNPTSYQLI
FSILSFKNRHLRLQWYTDLLRSPQTIIVDLLFGRRTIITANPENVEHILKTNFYN
PGKPFPTDLLGLGGIIFNSGELWSQKSLASHEFTWRSLEFTEILREEVQR
LIPVLSAIVDCGETVDFEVLKRFAPDVCKVSLGNDPDCDLTRPVPPELVKAFDVA
BISARRATEPVYAVKVKRVLNVGSEKRLREAIKTVHLSVSEIIRAKKSLDIGDVS
DKQDLRSFLAAGHEARVDSVIFSMAGRDTTSAAMTWLFWLLSQNDVDVETKILDE
LRNKSGLGFEDLREMSYTKACICEAMRLYPVAVMDSKHAANDDILPGCTPLKKGDK
VTYIPYGRMEKVGKWDDEFKRNREFEPEPSYGTXPVLKSVSSFKFPVQAGPRVC
IGKEMAFQMKYVGVSVLSRFLIIPVCNNRPVFPVPLTAHMAAGLKVKIKRRQCDSM
YI"
complement(join(25307..25416,25470..25577,25667..25731,
25857..25980,26228..26273,26940..27053))
/notes="gene id:MLE2.9
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08811.1"
/db_xref="GI:9758287"
/translation="WEASSSQPSDQTSKFLSDLPISSTVSSNPNGLRVY
ICEHDTSPQGLIKTQONLIRSLLLKQKGSSEKSKGTAEADCPKRAANRAJD
DRSSAKRAANASROEGSSRTERGDFQSLVEKRLAMEKGLPTKGRKATQSCTLN
SILKMLSGVASSSLFIVLPLIDYGT"
complement(27986..28738)
/notes="gene id:MLE2.10"
/codon_start=1
/evidence=not experimental
/product="transcription factor Hap5a-like protein"
/protein_id="BAB08812.1"
/db_xref="GI:9758288"
```



```
Query Match      6.4%; Score 277.8; DB 8; Length 58527;
Best Local Similarity 65.4%; Pred. No. 9e-46;
Matches 464; Conservative 0; Mismatches 222; Indels 23; Gaps 3;

QY 2792 AAAGGTAAGACATACAAATTTTGACATCCCTTTTACTAAATTTTAGTCCACCTGGC 2851
DB 5981 AATCTTGGATGGATGTTGATATGCTTTGTTTGGTGAATTCGAGGACCTCCAGG 5922

QY 2852 TCGCGAAGGAAACACAGTCACCGCTGATTAAAGATGAATTTTGTGTGGCCATTTAGCC 2911
DB 5921 TCAGGGAAGAGTACTCAATCTCCAGTAGTGAAGGATGATTGCTGTGTCACTTATCC 5862

QY 2912 ACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAAACCTCCACTGGGATTAAGGCTAA 2971
DB 5861 ACTGAGACATGTTAAGAGCTGCTGTGCTTCTTAAGACCCCTCTTGGTGTCAAGGCTAA 5802

QY 2972 GAAGCTATGGACAGAGTGTGTTTAAAGAAACATATAGCAACAGAAATTAATACCAGCAG 3031
DB 5801 GAAGCTATGAAAAAGTCAGTCTCTTTTGTCTTCTTGAGAGAGATTCTATTATTGTT 5742

QY 3032 AATGGGTTTCTTGATCTTTTGTGTTCTT-----TCCTTATCTTCTAGGAGAGCTTG 3083
DB 5741 CCTTTCGTTATAGATTTTATCTTCTTGTGATGTTTATTTTTCAGGGAGAGCTCG 5682

QY 3084 TTTCTGATGACTTGGTGTGTTGGGATTTATGATGAAGCCATGAAGAAACTTCATGCCAGA 3143
DB 5681 TCTCTGATGATTGGTGTGTTGTTATATGATGAAGCCATGAACAGCCAAATGTCAA 5622

QY 3144 AGGTTTATCTTGATGTTTCCCTAGAACTGTTGTTTCAAGCACAGAAAGGTGAGGTCT 3203
DB 5621 AAGGATTTATCTTGTATGGGTTCCCGAGGACTGTTACTCAGGCAGAGAGGTGATGCTCA 5562

QY 3204 TGGTCAATATGACCGCTATATAAAGAGCTCCTTTTGTGTTTATGAGCTGCTATATA 3263
DB 5561 ATAGAAAACTACTGTTGATC---GGTCTCTTTTGGTCTTTTGGTGTGTTGATGTTAA 5505

QY 3264 ATGACAGTTTCTATCATGTTATCATCTTTTCTTAAATAAATGTTGACGCTTGATGAA 3323
DB 5504 TTAATCTTCAACCATGTTGTTGTTCTTTGTT-----TGCAGCTCGATGAGA 5457

QY 3324 TGTGGCCAAACAGGTPACTTAAGATTGACAAGGTTCTTAATTTTGCATTTGATGATGCA 3383
DB 5456 TGCTTAAGAGCGAGGAACATGAATTTGACAAGTTCTCAACCTTGTATTTGATGATGACG 5397

QY 3384 TACTGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTAGATCTTATCATA 3443
DB 5396 TCTTGGAGGAAGAATAAACCGGGCGATGGATCCACCATCGAGTGCAGGAGTTACCACA 5337

QY 3444 CAAATTTGCTCTCTATGACTCCTGGACTTGATGATGATGAATGATCATAC 3492
DB 5336 CCAATTTGCTCTCTCCAAACCCCTGGAGTTGATGATGATGAAGACTTAC 5288

RESULT 4
RICKADKA
LOCUS      RICKADKA      1154 bp      mRNA      linear      PLN 02-FEB-1999
DEFINITION Oryza sativa mRNA for adenylate kinase-a, complete cds.
ACCESSION  D10334
VERSION    D10334.1  GI:391876
KEYWORDS   adenylate kinase-a; Adk-a.
SOURCE     Oryza sativa (strain:Yamahoushi) callus cDNA to mRNA,
           clone lib:lambda gt10 clone:Adk-a-29 and Adk-a-50.
ORGANISM   Oryza sativa
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
           1 (bases 1 to 1154)
AUTHORS    Kawai,M., Kidou,S., Kato,A. and Uchimiya,H.
TITLE      Molecular characterization of cDNA encoding for adenylate kinase of
           rice (Oryza sativa L.)
JOURNAL    Plant J. 2 (6), 845-854 (1992)
MEDLINE    93258424
REFERENCE  2 (bases 1 to 1154)

AUTHORS      Kawai,M.
TITLE        Direct Submission
JOURNAL      Submitted (30-JAN-1992) Maki Kawai, Univ. of Tokyo, Inst. of
              Molecular and Cellular Biosciences, Yayoi 1-1-1, Bunkyo-ku, Tokyo
              113, Japan (Tel:03-3812-2111(ex.7845), Fax:03-3812-2910)
FEATURES
  source
  1..1154
    /organism="Oryza sativa"
    /strain="Yamahoushi"
    /db_xref="taxon:4530"
    /clone="Adk-a-29 and Adk-a-50"
    /clone_lib="lambda gt10"
    /dev_stage="callus"
  gene
  1..1154
    /gene="Adk-a"
  CDS
  137..862
    /gene="Adk-a"
    /EC_number="2.7.4.3"
    /codon_start=1
    /product="adenylate kinase-a"
    /protein_id="BA01180.1"
    /db_xref="GI:391877"
    /translation="MAANLEDVPSMELTELLRRMKSSKPKRVLVGPCCGKGTQ
    SPLIKDFCLCHLATDMLRAAAVAAKTPLGTKAKEMDKBELVSDLVVGIIDAMKK
    TSCQGFILDGPPRTVVOAKLDEMLAQGTAKIDKVLNFAIDDALEERITGRWTHPS
    SGRSYHTKFAFPKTPGLDDVTGEPLIQKDDTAAVLKSRLFAFHVQTKPVIDYTKKG
    IVANLHAEKPKPEVTEVQKALS"
    polyA_signal
    1078..1086
    /gene="Adk-a"
  BASE COUNT  327 a  254 c  270 g  303 t
  ORIGIN
    Query Match      5.3%; Score 229; DB 8; Length 1154;
    Best Local Similarity 99.2%; Pred. No. 5.1e-36;
    Matches 241; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

    QY 1337 CAGAAACGAAACAAAAAGCCCAACACCTCTCGATCCGACCGAGACTCCTCCA 1396
    DB 1 CAGAAACGAAACAAAAAGCCCAACACCTCTCGATCCGACCGAGACTCCTCCA 60

    QY 1397 TTTTCAGGGCGGACGGGAGAGCAGCGAGTTCGTCACGACNAGGCTAGTGA 1456
    DB 61 TTTTCAGGGCGGACGGGAGAGCAGCGA--CGAGTTCGTCACGACNAGGCTAGTGA 118

    QY 1457 GTAGTTGTTCTCGGAGATGCGGCGAACCTGGAGGACGTCGCGTGGAGCTGATG 1516
    DB 119 GTAGTTGTTCTCGGAGATGCGGCGAACCTGGAGGACGTCGCGTGGAGCTGATG 178

    QY 1517 ACGGAGCTGCTCGCGCGATGAAGTGCAGCTCAAGCCCGACAGGGGTCATCCTCGTC 1576
    DB 179 ACGGAGCTGCTCGCGCGATGAAGTGCAGCTCAAGCCCGACAGGGGTCATCCTCGTC 238

    QY 1577 GGT 1579
    DB 239 GGT 241

    RESULT 5
    AB012248/c
    LOCUS      AB012248
    DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:MXI22.
    ACCESSION  AB012248 BA000015
    VERSION    AB012248.1  GI:3046856
    KEYWORDS
    SOURCE     Arabidopsis thaliana (strain:Columbia) DNA, clone lib:Mitsui P1
              clone:MXI22.
    ORGANISM   Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
              Rosidae; euroids II; Brassicales; Brassicaceae; Arabidopsi.
              1 (sites)
    REFERENCE Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Miyajima,N.
              and Tabata,S.
              and Tabata,S.
```


TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VI.
JOURNAL Sequence features of the regions of 1,367,185 bp covered by 19
MEDLINE Physically assigned P1 and TAC clones
REFERENCE DNA Res. 5 (3), 203-216 (1998)
AUTHORS 98403884
TITLE 2 (bases 1 to 80115)
JOURNAL Nakamura, Y.
 Submitted (23-Mar-1998) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research; 1332-3, Yana,
 Kisarazu, Chiba 252-0812, Japan (E-mail: ynakamura@kazusa.or.jp,
 Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MX122
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Grail-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://grenlin.zool.iastate.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is K6A12 and the 3' clone is M6A10.

FEATURES Location/Qualifiers
 1..80115
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MX122"
 /clone_lib="Mitsui Pi"
 join(5147..5278,5501..5587,5690..5758,5863..5995,
 6080..6126,6221..6307,6758..6819,7113..7155,7255..7335,
 7421..7534,7731..7836,7918..7981,8527..8596,8687..8848,
 8944..9072,9163..9312,9408..9494,9597..9665,9768..9866)
 /note="contains similarity to unknown protein
 Gb|AA049898.1
 gene_id:MX122.1"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB09450.1"
 /db_xref="GI:9758913"
 /translation="MGKTKYKPKGKTKTKAKADEKAKREGKKLSPEDDIDAIL
 LNIQEEAKKEHVHVEENVAAPSPRNSCSLTINPLKETELILYGFYNGQKTYVGD
 LRYDVEQEWKLVSSPPSPRSHQAQVAKNYLIYFGGFPTSPQERPHYKDFWM
 LDVKNQEWLNKGCPSRSGHRWLYKHKLIIFGGFYDLREVRVNDLYVFDLQ
 YKQWIKPKPGAMWPTARSGFFVQDEIFLYGYSKEVSEKSEKGIHVADLMSL
 DPTWENKVKIGMPSSRAGFCVCHKRALIFGVVDMEMEGDVMMSLFLNELYG
 FOLDNRWSDKSALQAGKTRLDQVSVSEVVKPCGRINCMVVGKOTLYTYGGMMG
 IKDKVTLDLYLSNLKLDWKCIIPITTEWVSDDEDEGDDDDSDSEGCNSE
 ESDDDDEEVEAMVDGSKVGVVAVMIKQGSLSRREKRAIEQIRANGLSDSO
 RPEVGEETLKPKYKTNMYQWAAVEHTQHTGKELRKGDFDLAEITYLELPIDELTA
 ILEAQKAEAEAEPPASSSSRRGGIAKKR"
 10440..11115
 /note="contains similarity to mutator-like transposase
 gene_id:MX122.2"
 /pseudo
 /codon_start=1
 /evidence=not experimental
 complement(join(11785..11943,12033..12185,12697..12801,
 12893..13132,13192..13358,13545..13719,14023..14195,
 14294..14363,14921..15378))
 /note="gene_id:MX122.3"
 /codon_start=1

/evidence=not experimental
 /product="histone acetyltransferase"
 /protein_id="BAB09451.1"
 /db_xref="GI:9758914"
 /translation="MNGELKKPRPGKGGYQGRGLTEEARVRAISEIVSMIERSHR
 NENVDLNAIKTACRKYGLARAPKLVEIMIAALPDSERETLLPKRAKRVNPNVQA
 AVMGKPHRCPIATTTGNIQVYCPGPDSDPSTQSYGTEPTSMRAIRARVNPVQA
 RSRIDKRLGLHSDVKVEFTLMGCTFMSLPAEYRDFFINLHDALSGHTSANVEAVA
 YSEHSATKCIQMTIECPDYCLGPHLRQMLYCTGLEIGVOSTVEDVARDTGKIVTLV
 AAVADCCAKADAGFKVAHMPDLPNGVVERMESFEFPSPSFRADGKIVTLV
 IRGELTLMKIRYINPEQLVDIVAKILSMVPTWYRVQRDIIPMLVTSVEK
 GNRLRLAARDMDLGLKCRDVTREAGIQIHHKIKRQVRLVRLVAVTVAGVHGRDADKLQHG
 YGTLMREARIRARRHRSNKIGVSPMGKTSVRELHVYGTAVPVMHGRDADKLQHG
 complement(join(16097..16288,16390..16543,17220..17342,
 17523..17633,17815..17939,18018..18146,18527..18721,
 18809..18937,19035..19148))
 /note="contains similarity to ABC transporter
 gene_id:MX122.4"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB09452.1"
 /db_xref="GI:9758915"
 /translation="MHGFDFKEIQEKISDSFRPQWRSFQVWRATNIVTYGVFOLRV
 SLVDKAKQEMMERHOEAADKIYFMCSDLGDFLKIQAOLAKPDMAPAAWVKLVLT
 LCDQAPTPPDALQIVLEKLGKSGIEIPETDEKPLGSASIAQVQHPIELRMWTDI
 ENLQALFYQWRDIDKEDLHSTKENEKQIGYEFDEKREANAMEIRCFELNNKSP
 VIVPRVDRWTKVLYWETNGIPILSYGDEMAKKGINPHGKIAEAAKHNLNLSLR
 AYGMILKSGFFHADPHPNILICKQEAHSHILEMGLHTVACKEQVQLKLAQTLLF
 DTQMTQTVLQVFPSSDIKKIAVTFPEELFSVLRTVTLRGLSGVNGVNSCAEQ
 WRAMEALASGRVTRGKWLRYVLTTLFA"
 complement(join(22220..22387,22459..22531,22650..22732,
 22903..22941,23062..23139,23402..23557,23788..23853,
 24049..24129,24218..24280,24552..24659,24908..24991,
 25355..25405,25557..25615,25689..25744,25838..26127))
 /note="gene_id:MX122.5"
 /codon_start=1
 /evidence=not experimental
 /product="DNA repair protein-like"
 /protein_id="BAB09453.1"
 /db_xref="GI:9758916"
 /translation="MNKSKGKKTVMVCECSHSGSQWMSGRACHKQVMTKRSEF
 SSSASGGNGTGLGTEGKGLSWLPEQATVQPHRLTDVIGHITQQQWRISLPLGLFG
 NEVARVLGGELAPGLJLIGDPIGKSLILQGSLELAEPAPVLVISEEVSQIGSR
 ADNRVITTEELYLFSSDLQDILINKAHLSPRALIIDSITQVYLDKDVTSAGGLTQVK
 ECTSTLLRAKKNIPVFLGVHTKAGDIAGPVLEHIVDVVLYLEGEBSYRLRLS
 VNRFGSDDELGVFEMANSLEVYVSNPISQNPDSVLAAGLAVAVMDGSKSL
 IEVQALCPGSTVSRHVGQASRADMIAVLMKQAGLRIQENGFVANGMALSET
 AGDLAIAAICSSFLPPIPHGVAFIGEIGLAGEVTRVPRMEKRVSTVAKLGFNKC
 PKSVESLKLALKEIIEIGKNLKLINAVPKG"
 complement(join(29282..29509,29616..29945,30038..30175,
 30265..30384,30548..30746,30938..31105,31525..32196))
 /note="gene_id:MX122.6"
 /unknown protein
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB09454.1"
 /db_xref="GI:9758917"
 /translation="MATSAPNTMKRTTSLAKSESQDSSYSTRSARRHRSLSRF
 SHRPDDPIEVTNMRKGVNTVRGSGFGISLDDLAVEFPSPFSGESEISGSGRGRS
 CLRWSGGGTGVNTNSORGRSVRVSGGGGGLRLDADTSSRRSRLSPASN
 NRGSSKVDVSVNNSSRRRSVSRQPNRVERKVENSGVIEGDRENANSSRRSLSVRR
 RIENSESDVDDQEGSSSRDVRVFGMSKQNSGQSAASDNKQGLRSFSQNP1KYH
 DYGSSQAVTDDQEGSSSSKGTIRIINTVYQNKATPKKRSGLNSGYSQKRS
 DNRGVTSTKGYATKLQESERRELLAIMEEQRELSELVNLKEILTENSBAEE
 KPLRTRKSKDRSRMSKCLTDEAQFIDEFISNIEDTDFSLSEDESSSSFGMIK
 TQSSQTSALKSTSVEMDGVMLQWETPDVSAVAACILKSPHTPNKSLVWESDIT
 QDASGRGTSIGTSRRSGSPYDVKPWPTKLIKIDVAEVLKRPNSDILNETWK
 LERHTSSGSLVLCSSRLI"
 complement(36018..36893)
 /note="emb|CAB62340.1
 gene_id:MX122.7
 similar to unknown protein"

```
/codon_start=1
/evidence=not experimental
/protein_id="BAB09455.1"
/db_xref="GI:9758918"
/translation="MNFPGGCCIRYGGSGGNDMSKVDRLMLRVPRIAPRDPDSGGSP
ASPTKNGSVITNTYSSRRKGRKYSKENNSSSTGSVNGNSKQRNDETKNGSGGG
REIVTLPLLPETPKDQPLKAKAALPGLAALWLSFNDGASYNRYQTELMTETWS
SLLTVECTERLMEGEYELGCTDEERKMNLERDTCPGFISDGLGRVITNGSYRLVV
GKDHQCSKMSVLWMEKEPLVIRYRFTCRMRLQYTCRDKEYSSITSFCDVWRMSDGG
FAWLDVDAALCLGR"
complement(join(39967..40068,40144..40233,40331..40498,
40586..40851,41096..41216))
/note="gene_id:MX122.8"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB09456.1"
/db_xref="GI:9758919"
/translation="MATSSAASVDMEDIQTVLMSLELLRMKCSKPKRLVFIQPPG
SGKGTQSPVINDFELCHLSTGDMKRAAANKPTLGVKAKEAMDKGLVSDLLVVGIM
DEANRPKQKGFILDGFPRTVTOAEKLDEMLNRGAQIDKVLNFAIDDSVLEERTG
RWIHPSSGRSYHTKFAFPKGVDDLTGTEPLIQKDDNADVLRLSLDAFHQKQTPVID
YYAKKENLVNPAEKAPPEVTKVKKVST"
join(42526..42663,42914..42997,43149..43190,43341..43451,
43540..43731,43935..44032,44339..44516)
/note="gene_id:MX122.9
unknown protein"
/pseudo
/codon_start=1
/evidence=not experimental
complement(46967..48880)
/note="gb|AAD25781.1
gene_id:MX122.10
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB09457.1"
/db_xref="GI:9758920"
/translation="MLLIFFSFDNRLSNIKNTAMTNOEONQNDALVARLEAAESVIHRW
DGNDSRRHSSSGNRYRSSFSLSFDESPEATEFLSAVDEIISLLEDLSSENKPD
Query Match 3.9%; Score 171.2; DB 8; Length 80115;
Best Local Similarity 59.9%; Pred. No. 41e-24;
Matches 400; Conservative 0; Mismatches 148; Indels 120; Gaps 2;
QY 2827 TTAATAATTTAGTCCACCTGGCTCGGAAAGGGAACACAGTCCGCTGATTAAGGA 2886
DB 40865 TTTGGAATTTGCAGGTCCACCTGGTTTCAGGGAAGGTACACAGTCTCCGGTCTATAAGGA 40806
QY 2887 TGAATTTTGTGCTGCCATTTAGCCACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAA 2946
DB 40805 TGAGTTTTTGTGTGTCATTTGTCTACTGGTGTGACATGCTTAGAGTCTGCTGCTGCTAA 40746
QY 2947 AACTCCACTTGGGATTAAGGCTTAAGAAGCTATATGGAACAAGTAGTTTTTAAAGAACATAT 3006
DB 40745 GACTCCTCTGTGTGTAAGGCAAGGAAGCAATGGATA----- 40708
QY 3007 AGCAACAGAAATATAACACGAGGAGTGGTTCTTGTGATTCCTTTGTTCTTCCTTAT 3066
DB 40707 ----- 40708
QY 3067 CTTTAGGAGAGCTGTTTCTGATGACTTGGTTGTTGGGATTTATGATGAAGCCATGAA 3126
DB 40707 -----AGGAGAGCTGTTTCTGATGACTTGGTTGTTGGTATCATGTGATGAAGCAATGAA 40653
QY 3127 GAAAACTTCATGCCAGAAAGTTTATCCTTGATGGTTTCCCTAGAACTGCTGTTCAAGC 3186
DB 40652 CAGACCCAAATGTGAGAAGGTTTCATTTGATGGGTTCCCGGAGCTGTGACTCAAGC 40593
QY 3187 ACAGAAGGTGAGTCTCTGGTCAATATGACCCGCTATATAAAGAGAGCTCCTCTTTTGTAT 3246
DB 40592 AGAGAAGGTTTCTGTT-----CCTTTCTGTG 40566
```

```
QY 3247 TAGAGCTGCTATATAAATGCACAGTTTCTATCATTTGTTATCACTTTTCTTACTAAAAAAT 3306
DB 40565 GTTAGTTGTTTGTGATCTCTCTGCTCTTGTAAAGTTCTGACTTTTGTGTTTATAAT 40506
QY 3307 GGTGAGCTTGATGAATGTTGGCCAAACAGGTACTAAGATTGACAAGTTCTTAAATTT 3366
DB 40505 CTTGTAGCTTGATGAGATGCTTAAATAGAGGGGAGCTCAGATAGATAAGGTGCTTAAATTT 40446
QY 3367 TGCATTTGATGATGCAATATCTGGAAGAACGATTACCGGCTCGTTGGATCCACCCATCAAG 3426
DB 40445 TGCATCGATGATTCGGTTCTCGAAGAAGAATTAATGGAAGTGGATTCACCTTCAAG 40386
QY 3427 TGGTAGATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGGACTTGGATGATGAAG 3486
DB 40385 TGGAAGAAGCTATCATACTAAATTCGCACCTCTCTAAAGTTCAGGAGTCCGATGATGAAG 40326
QY 3487 TCATACCA 3494
DB 40325 AAAAACCA 40318
RESULT 6
RICKN57
LOCUS RICKN57 Rice mRNA for adenylate kinase (KN57 gene), linear PLN 02-FEB-1999
DEFINITION Rice mRNA for adenylate kinase (KN57 gene), partial sequence.
ACCESSION D10429
VERSION D10429.1 GI:218170
KEYWORDS adenylate kinase.
SOURCE Oryza sativa
ORGANISM Oryza sativa callus cDNA to mRNA.
REFERENCE 1. (sites)
AUTHORS Uchimiya,H., Kidou,S., Shimazaki,T., Aotsuka,S., Takamatsu,S.,
Nishi,R., Hashimoto,H., Matsubayashi,Y., Kidou,N., Umeda,M. and
Kato,A.
Random sequencing of cDNA libraries reveals a variety of expressed
genes in cultured cells of rice (Oryza sativa L.)
Plant J. 2, 1005-1009 (1992)
2. (bases 1 to 387)
Kidou,S., Umeda,M., Kato,A. and Uchimiya,H.
Plant cDNA homologue to rat insulinoma gene encoding ribosomal
protein S15
Nucleic Acids Res. 21 (8), 2013 (1993)
93261842
8493114
REFERENCE 3. (bases 1 to 387)
Uchimiya,H.
Direct Submission
Submitted (27-JAN-1992) Hirofumi Uchimiya, Institute of Mol. &
Cell. Bioscience, The University of Tokyo, Department of Cellular
Function; 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
E-mail:huchimiya@tansei.cc.u-tokyo.ac.jp
Tel:03-3812-2111(ex.7844), Fax:03-3812-2910
Submitted (27-Jan-1992) to DDBJ by:
Hirofumi Uchimiya
Institute of Applied Microbiology
University of Tokyo
Bunkyo-ku, Tokyo 113
Japan
Phone: 03-3812-2111 x7844
Fax: 03-3812-2911.
Location/Qualifiers
1. .387
/organism="Oryza sativa"
/db_xref="taxon:4530"
/tissue_type="callus"
BASE COUNT 108 a 71 c 104 g 104 t
ORIGIN
Query Match 3.8%; Score 164.4; DB 8; Length 387;
Best Local Similarity 74.0%; Pred. No. 6e-23;
```



```

gene      join(37075..38148,39504..39906,40314..40415,40513..40736)
          /gene="P0698A10.11"
CDS       join(37075..38148,39504..39906,40314..40415,40513..40736)
          /gene="P0698A10.11"
          /note="contains ESTs AU101336(E11055), AU101337(E11055)"
          /codon_start=1

Query Match      2_start; Score 115.6; DB 8; Length 137348;
Best Local Similarity 76.3%; Pred. No. 8.3e-13;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 893 TTTCAACGATCATATAACGAGAAAGCTCATTAGACATTTATTCTTATGATATTTTATAAT 952
      |||||
Db 136817 TTTTAACTCACCACCAACAGAAAACCTTATTAGCACATAATTAATTAAGTATTAAT 136758

QY 953 TATAAATTTGAAAAAATATTTTATTGAAATTTTTTAAACAAATGTATGCATAAATTTT 1012
      |||||
Db 136757 TATAAATTTGAAAAAATATTTTATTGATTTTTTAAATAAATTTTATATAGATTTTTT 136698

QY 1013 TTTAAACACACCAATTTAAACCTTTTAAAGCATCTCTAATAGGAACGAGGAAGTTAAA 1072
      |||||
Db 136697 TTTAAATAACATATCATTTAAACAGTTTGAAGAAAGCATCTAACGGAACGATAAAATAAA 136638

QY 1073 GATTCA 1078
Db 136637 GTTTGA 136632

RESULT 9
AY079340
LOCUS      Arabidopsis thaliana putative adenylylate kinase (At5g63400) mRNA,
           complete cds.
ACCESSION      AY079340.1 GI:19310680
VERSION        AY079340.1
KEYWORDS       FLI CDNA.
SOURCE         Arabidopsis thaliana.
ORGANISM       Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 772)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Egtu,P.,
Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H.,
Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
Theologis,A. Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 772)
Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT        The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
Hayashizaki,Y. and Shinozaki,K.

TITLE          The Salk, Stanford, PGEC (SSP) Consortium members constructed and
JOURNAL        sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
AUTHORS        Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
               Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
               Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
               Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
               Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

```

```

Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.

FEATURES             Location/Qualifiers
     source            1..772
                        /organism="Arabidopsis thaliana"
                        /db_xref="taxon:3702"
                        /chromosome="5"
                        /clone="U09821"
                        /note="This clone is in PUNI 51."
     ecotype            Columbia"
     gene              1..772
                        /gene="At5g63400"
     CDS                1..741
                        /gene="At5g63400"
                        /codon_start=1
                        /evidence=experimental
                        /product="putative adenylylate kinase"
                        /protein_id="AA185071.1"
                        /db_xref="GI:19310681"
                        /translation="MATGGAADLEDVQVTLMSLLRLKCSQPKRLIFIGPPGS
                        GKGTSVPVKDEYCLHLSGTGMLRAAVASPTPLGVKAKEAMEKGLVDDLVVGIIID
                        EAMNPKCKGKGFILDPRTVTOAEKLEMLKRRGTEDKVLNFAIDAIEERITGR
                        WHPSSGSYHTKFPAPKPTGVDITGEPLIQKDDNADVLKSLAAFHQSOTQPVIDY
                        YAKKAVLTNIQAEKAPQEVTSVKALS"
                        742..772
                        /gene="At5g63400"
     3'UTR             221 a 175 c 189 g 187 t
     BASE COUNT        221 a 175 c 189 g 187 t
     ORIGIN             2.6%; Score 112.8; DB 8; Length 772;
                        Best Local Similarity 76.7%; Pred. No. 1.9e-12;
                        Matches 138; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 3308 GTGCAGCTTCATCAAAATGTTGCCAAACAAAGGTACTTAAGATTGACAAAGTTCTTAAATTTT 3367
      |||||
Db 379 GAGAAGCTCGATGAGATGCTTAAGAGCGGAGGAAGTGAATTTGACAAAGTTCTCAACTTT 438

QY 3368 GCAATTGATGATGCATATCTAGGAAGAACCAATACCGTTCGTGGATCCACCATCAAGT 3427
      |||||
Db 439 GCTATTGATGACCAATCTTTGGAGGAAGAATAACCGGCGATGGATCCACCATCGAGT 498

QY 3428 GGTAGATCTTATCATACAAAATTTGCTCTCTCTTAAGACTCTCGGACTTGATGATGTAAGT 3487
      |||||
Db 499 GGCAGGAGTTACCACACCAAAATTTGCTCTCTCCAAACCCCTGGAGTTGATGATATTACT 558

RESULT 10
AY085188
LOCUS      Arabidopsis thaliana clone 13667 mRNA, complete sequence.
DEFINITION      Arabidopsis thaliana clone 13667 mRNA, complete sequence.
ACCESSION      AY085188.1 GI:21403898
VERSION        AY085188.1
KEYWORDS       FLI CDNA.
SOURCE         Arabidopsis thaliana
ORGANISM       Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 904)
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
2 (bases 1 to 904)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished

```


Db 393 GAGAACTCGATGAGATGCTTAAGAGCGGAGGAACTGAAATTCACAAAGTTCTCAACTTT 452
 QY 3368 GCAATTGATGATCAATCTACTGGAGAACGAATTTACGGTTCGTGGATCCACCATCAAGT 3427
 Db 453 GCTATTGATGACCAATCTTTGGAGGAAAGAAATAACCGGCGGATGGATCCACCATCGAGT 512
 QY 3428 GGTAGATCTTATCATCAAAATTTGCTCTCTCTAGACTCCTGGACTTGATGATCAAGT 3487
 Db 513 GGCAGGAGTTACCAACCAAAATTTGCTCTCTCCAAAACCCCTGGAGTTGATGATATTACT 572

RESULT 12
 LOCUS AP005594 124815 bp DNA linear HTG 24-JUL-2002
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone P0701E06,
 *** SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AP005594
 VERSION AP005594.1 GI:21952963
 KEYWORDS HTG: HTGS PHASE2.

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:P0701E06.

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T., Hattori,M., Sasaki,Y. and Katayose,Y.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC

clone:P0701E06

PUBLISHED ONLY In Database (2002)

REFERENCE 2 (bases 1 to 124815)

AUTHORS Sasaki,T., Matsumoto,T., Hattori,M., Sasaki,Y. and Katayose,Y.

TITLE Direct Submission

JOURNAL Submitted (23-JUL-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source

1. .124815

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="9"

/clone="P0701E06"

BASE COUNT 37055 a 25766 c 25390 g 36503 t 101 others

ORIGIN

Query Match 2.5%; Score 109.2; DB 2; Length 124815;
 Best Local Similarity 63.2%; Pred. No. 1.6e-11;
 Matches 168; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 820 TAAGACGATGAGTTGGAAACAGAGAAACACAGCGCTTAAGGCTTTGTGATCTCTTA 879

Db 56107 TTATAAGTATGTTGTAGGATTAATCAATTAAGATGACTAATGCTCGTTCATACCTC 56166

QY 880 GTTCGAGTTGATTTTCAACGCGATGATAACGAGAAAGCTCAATAGCACATTATTACTT 939

Db 56167 ATTTGGAGATGAGAATTTCAACGCGCACAAAACGAGAAATCTCATAGCACATGATCAAT 56226

QY 940 AGATATTTATTAATTAATAAATTTGAAAAAATATTTATTTGAAATTTTAAACATGATG 999

Db 56227 AAGTGTTAATTAATTAATAAATTTGAAAAATGGAATTTATTTGATATTTTAACTACTTCTA 56286

QY 1000 CATAAATTATTTTTTAAAAACACACCAATTTAACCCCTTTAAAAAGCATCTCTAATAGAAA 1059
 Db 56287 TATAGATTTTTTTTTTAAAGAACACATCATTTAACTGTGTGAAAAAGTATGCTAATGAAAA 56346

QY 1060 CGAGGAAGTTAAAGATTACACCGAAGT 1085

Db 56347 TGAGAGAAGTTGAAGTTTACTAGAGT 56372

RESULT 13

LOCUS AP003261/c

DEFINITION

Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,

PAC clone:P0471B04.

ACCESSION AP003261

VERSION AP003261.3 GI:22202658

KEYWORDS

SOURCE

Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:P0471B04.

ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC

clone:P0471B04

PUBLISHED ONLY In Database (2001)

REFERENCE 2 (bases 1 to 138025)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT On Aug 9, 2002 this sequence version replaced gi:21202837.
 Genes were predicted from the integrated results of the following:
 GENSCAN1.0, BLASTX2.0, BLASTY2.0 as well as SplicePredictor
 (October 1998 version). The genomic sequence was searched against
 NCBI NonRedundant Protein database, nr
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
 RGP. Protein homologies of the coding regions were searched against
 NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
 the identified cDNA sequences using BLASTN 2.0 with the
 corresponding DBJ accession no. and RGP clone ID.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with EST homology (covering
 almost the entire length of partial sequence) is classified as an
 'unknown' protein. A gene predicted with a gene prediction program
 is classified as a 'hypothetical' protein.
 The orientation of the sequence is from SP6 to T7 of the PAC clone.
 This sequence of P0471B04 clone has an overlap with P0698A10 clone
 (DBJ : AP003297) at the position 1 to 24,438 of 5' end and an
 overlap with P0018C10 clone (DBJ: AP003227) at the position 48,801
 to 138,025 of 3' end. Detailed information on overlap and assembly
 quality together with annotation of this entry is available at
 http://rgp.dna.affrc.go.jp/GenomeSeq.html.

FEATURES

source

1. .138025

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="1"

/clone="P0471B04"

join(9533..9715,9887..10019,10850..10988,11365..11401,
 13640..13931,14823..14955,15564..15576)

/genes="P0471B04.1"

join(9533..9715,9887..10019,10850..10988,11365..11401,
 13640..13931,14823..14955,15564..15576)

gene

CDS


```

Db 23907 TTTTAACTCAGCAAAACAGAAAACTTATTAGCACATAAATTAATTAAGTATTAAAT 23848
Qy 953 TATAAAGCTGAAAAAATAATTTATTGAAATTTTTTAAACAATGTATGCNAATAATTTT 1012
Db 23847 TATAAAGCTGAAAAAATAATTTATTGAAATTTTTTAAACAATGTATGCNAATAATTTT 23788
Qy 1013 TTAACAACACACCAATTTAAACCTTTTAAAGAGCATCTTAATAGGAACGAGGAAGTTAAA 1072
Db 23787 TTAATAACATATCATTTTAACGTTTGAAGAAGCATCTAACGAAACGATAAAATTTAAA 23728
Qy 1073 GATTCA 1078
Db 23727 GTTTGA 23722

RESULT 14
CNS07YDPX
LOCUS
DEFINITION
Oryza sativa chromosome 12 clone OJ1618_C05, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION
AL713952
VERSION
AL713952.2 GI:22003292
KEYWORDS
HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE
Oryza sativa
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 169663)
AUTHORS
Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Segurens,B., Pellatier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 169663)
Genoscope.
Direct Submission
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jul 29, 2002 this sequence version replaced gi:19715885.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequencing may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source
1. 169663
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosomes="12"
/clone="OJ1618_C05"
BASE COUNT 46256 a 37378 c 36384 g 49645 t
ORIGIN
Query Match 2.5%; Score 109.2; DB 2; Length 169663;
Best Local Similarity 68.8%; Pred. No. 1.7e-11;
Matches 150; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
Qy 861 AAGCTCTTCATCCTCTAGTCGAGGTGATTTTCAACGCGATGATAACGAGAAAGCT 920
Db 103185 AAGCACTAGGCTTCGGCTTGTGATAGAAAAATTTTAAACGACGCAAAACAGAAAGCT 103244

```

```

Qy 921 CATTAGCACATTTATTACTTAGATATTTATTAATAATAAACTTCAAAAAATATTTTGA 980
Db 103245 CATTAGCACATAATGATTAAAGTATTAAATTAATAATAAACTTCAAAAAATGATTTTGA 103304
Qy 981 ATTTTTTAAACAATGATGCATAAAATTTATTTTAAACAACACACCAATTTAAACCCCTTTAA 1040
Db 103305 TTTTTTAAACAATCTCTATATAGAAATTTTCTTAAACAACACATCATTTAAACAGTTTAA 103364
Qy 1041 AAGCATCTCTAATAGAAACGAGGAAGTTAAAGATTCA 1078
Db 103365 AAAGCGTGCTTAACAGAAAAACGAGGAAGTTGAAAGTTTAA 103402

RESULT 15
OSJN00015
LOCUS
DEFINITION
Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB0079B02,
complete sequence.
ACCESSION
AL606455
VERSION
AL606455.2 GI:21740404
KEYWORDS
HTG
SOURCE
Oryza sativa.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1
AUTHORS
Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L.,
Liu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B.,
Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Fu,G.,
Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F.,
Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang W., Wang,L.J.,
Ding,C.W., Sheng,H.H., Gu,Q.L., Chen,S.T., Ni,L., Zhu,F.H. and
Hong,G.F.
Direct Submission
Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB0079B02.
On Jul 12, 2002 this sequence version replaced gi:15552693.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
genscan (http://cgr-081.mit.edu/GENSCAN.html), GeneMarkHM
(http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
complete sequence against NCBI none redundant protein database (nr)
(http://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES
Location/Qualifiers
source
1. 126659
/organism="Oryza sativa"
/variety="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosomes="4"
/clone="OSJNB0079B02"
BASE COUNT 35469 a 27283 c 28042 g 35865 t
ORIGIN
Query Match 2.5%; Score 108.8; DB 8; Length 126659;
Best Local Similarity 55.9%; Pred. No. 2e-11;
Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

```


GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:09:00 ; Search time 809.479 Seconds
(without alignments)
12076.811 Million cell updates/sec

Title: US-09-802-937-1

Perfect score: 4341

Sequence: 1 cgcgcgaagattaattagg.....gtatatcttcattgtttt 4341

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------|
| 3 | 1440 | 33.2 | 1440 | 21 | AAA07482 |
| 4 | 1440 | 33.2 | 1440 | 21 | AAA10961 |
| 5 | 115.8 | 2.7 | 809 | 21 | AAC48619 |
| 6 | 115.2 | 2.7 | 482 | 21 | AAC48110 |
| 7 | 113.2 | 2.6 | 459 | 21 | AAC38309 |
| 8 | 112.8 | 2.6 | 904 | 21 | AAC47002 |
| c 9 | 112.8 | 2.6 | 981 | 24 | ABN98503 |

| | | | | | |
|------|-------|-----|-------|----|----------|
| 10 | 102.6 | 2.4 | 608 | 21 | AAC49033 |
| 11 | 102.6 | 2.4 | 611 | 21 | AAC35687 |
| 12 | 102.6 | 2.4 | 1017 | 21 | AAC49022 |
| 13 | 102.6 | 2.4 | 1023 | 21 | AAC37085 |
| 14 | 102.2 | 2.4 | 295 | 24 | ABL72417 |
| 15 | 101.8 | 2.3 | 894 | 24 | ABN98716 |
| 16 | 100 | 2.3 | 422 | 24 | ABL49953 |
| 17 | 97.4 | 2.2 | 423 | 24 | ABL49955 |
| 18 | 97.4 | 2.2 | 424 | 24 | ABL49954 |
| c 19 | 66.6 | 1.5 | 5875 | 24 | ABL32289 |
| c 20 | 64 | 1.5 | 6048 | 24 | ABO67002 |
| c 21 | 64 | 1.5 | 9642 | 24 | ABL32357 |
| c 22 | 63 | 1.5 | 6095 | 22 | AS46310 |
| c 23 | 63 | 1.5 | 6095 | 24 | ABL70150 |
| c 24 | 63 | 1.5 | 6095 | 24 | ABL32361 |
| c 25 | 63 | 1.5 | 6095 | 24 | ABL34475 |
| c 26 | 62.6 | 1.4 | 8237 | 22 | AS46802 |
| c 27 | 62 | 1.4 | 721 | 21 | AAF14639 |
| c 28 | 61.6 | 1.4 | 645 | 23 | AS453274 |
| c 29 | 61.4 | 1.4 | 682 | 21 | AS44208 |
| c 30 | 61.4 | 1.4 | 883 | 22 | ABL15210 |
| c 31 | 61.4 | 1.4 | 15548 | 24 | ABL34155 |
| c 32 | 61.2 | 1.4 | 7165 | 24 | ABL70222 |
| c 33 | 61.2 | 1.4 | 7165 | 24 | ABL32751 |
| c 34 | 61.2 | 1.4 | 7165 | 24 | ABK31267 |
| c 35 | 61 | 1.4 | 736 | 21 | AAF07686 |
| c 36 | 60.8 | 1.4 | 3687 | 23 | ABL05676 |
| c 37 | 60.6 | 1.4 | 38342 | 22 | AS46746 |
| c 38 | 60.6 | 1.4 | 38342 | 24 | ABK31507 |
| c 39 | 60.4 | 1.4 | 6079 | 24 | ABL32421 |
| c 40 | 60.4 | 1.4 | 7442 | 22 | AS46686 |
| c 41 | 59 | 1.4 | 38342 | 22 | AS46746 |
| c 42 | 59 | 1.4 | 38342 | 24 | ABK31507 |
| c 43 | 58.8 | 1.4 | 13202 | 24 | ABL33485 |
| c 44 | 58.6 | 1.3 | 11836 | 22 | AS45395 |
| c 45 | 58.6 | 1.3 | 11836 | 24 | ABK28240 |

ALIGNMENTS

RESULT 1

AAA07481

ID AAA07481 standard; DNA; 4341 BP.

XX

AC AAA07481;

XX

DT 17-JUL-2000 (first entry)

XX

DE Rice adenylate kinase DNA sequence fragment #1.

XX

XX Rice; adenylate kinase; promoter; structural gene expression;

KW plant gene expression; herbicide resistance; stress resistance;

KW light resistance; antibacterial production; flowering induction;

KW insecticidal substance production; ss.

XX

OS Oryza sativa.

XX

PN WO200015812-A1.

XX

PD 23-MAR-2000.

XX

PF 08-SEP-1999; 99WO-JP04847.

XX

PR 10-SEP-1998; 98WO-JP04088.

XX

XX (NISC) NISSAN CHEM IND LTD.

PA

XX Rice adenylate kin

XX Partial rice adeny

PI Arabidopsis thalia

XX Zea mays DNA fragm

XX Zea mays DNA fragm

DR Arabidopsis thalia

XX Arabidopsis thalia

PT Rice adenylate kinase gene-originated DNA fragment with promoter

PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
XX induction in recombinant plants -

Claim 1; Page 25-27; 33pp; Japanese.

XX This sequence represents a fragment of the rice adenylate kinase gene.
CC The invention relates to a DNA fragment (such as this sequence), that has
CC a regulatory function (preferably acting as a promoter), that has
CC desired foreign structural genes in a plant when introduced into a
CC plant body or plant cell. The DNA fragment can be inserted into a plant
CC to control the expression of desired foreign structural genes, useful for
CC inducing in the plant e.g. herbicide, stress and light resistance,
CC production of antibacterial and insecticidal substances, sucrose
CC induction in recombinant plants, and in promoting flowering and crop
CC yield. The DNA fragment has a regulatory function on the expression of
CC desired structural genes in a plant, particularly with control of the
CC expression in specific tissues and locations to provide safe chemicals.

XX Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;

Query Match 100.0%; Score 4341; DB 21; Length 4341;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60

DB 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60

QY 61 AATCAATCGGGTGTTCGTTGGACATGTTTTCGAGGAAATTAAGCAAGAAATTAAG 120

DB 61 AATCAATCGGGTGTTCGTTGGACATGTTTTCGAGGAAATTAAGCAAGAAATTAAG 120

QY 121 AAGAACTCTCAAGCTGACATGAGAAACGTAATCCAAATGGAACGGAATTCAGTCGTC 180

DB 121 AAGAACTCTCAAGCTGACATGAGAAACGTAATCCAAATGGAACGGAATTCAGTCGTC 180

QY 181 TCTTGTAATCACTGTTTGTAGATACATAGACAGTCCCAACGTTTGATGGCTCCTATTGG 240

DB 181 TCTTGTAATCACTGTTTGTAGATACATAGACAGTCCCAACGTTTGATGGCTCCTATTGG 240

QY 241 CTGCTGTGATGACTGCTGTGTCACAAAGCATCAAAATGCTTCTGGAGTATCTTTATTA 300

DB 241 CTGCTGTGATGACTGCTGTGTCACAAAGCATCAAAATGCTTCTGGAGTATCTTTATTA 300

QY 301 CCGAAAAACCCCAAGATTATCTTATCCACTCAGGGTAAATGTGCTGAACATATGCAATG 360

DB 301 CCGAAAAACCCCAAGATTATCTTATCCACTCAGGGTAAATGTGCTGAACATATGCAATG 360

QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTGCTCAAAATGGAATTTGAGTCCAA 420

DB 361 AATACAAATTCGCAAAATATCATGTTATCTATCTGCTCAAAATGGAATTTGAGTCCAA 420

QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTAATTAATTTTTCATGAAA 480

DB 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTAATTAATTTTTCATGAAA 480

QY 481 CGCAATTCACCGTTCGAGAAATATGCTGTCAATAAATAGTAGTCTAGTGCAGAAACAAA 540

DB 481 CGCAATTCACCGTTCGAGAAATATGCTGTCAATAAATAGTAGTCTAGTGCAGAAACAAA 540

QY 541 ATTAATATACATAAAAGAGGTTGTTAAATACAAACCATGTTTCGTAACACTCT 600

DB 541 ATTAATATACATAAAAGAGGTTGTTAAATACAAACCATGTTTCGTAACACTCT 600

QY 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCTATTTTCCAATTAAGAAAAATAAAGCT 660

DB 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCTATTTTCCAATTAAGAAAAATAAAGCT 660

QY 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAGAAAAATA 720

DB 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAGAAAAATA 720

QY 721 TGATGGTTTATTAATATGATTTTTTTTTTAAATAAATCTTTACATATAATTTCTTTAGGAAACA 780

DB 721 TGATGGTTTATTAATATGATTTTTTTTTTAAATAAATCTTTACATATAATTTCTTTAGGAAACA 780

QY 781 TATCATTTTAAATGTTTGAATAACGTCACATATAGAAACCTTAAGAACGATGAGTTGGGAAA 840

DB 781 TATCATTTTAAATGTTTGAATAACGTCACATATAGAAACCTTAAGAACGATGAGTTGGGAAA 840

QY 841 CAAGAGAAAAACACAGCCTTTAAGCCTTTTGTATCTCTAGTTGGAGGTTGATTTTCAAAAC 900

DB 841 CAAGAGAAAAACACAGCCTTTAAGCCTTTTGTATCTCTAGTTGGAGGTTGATTTTCAAAAC 900

QY 901 GCATGATAAACGAGAAAGCTCATAGACATATTAATTAATTAATTAATTAATTAATTAATTA 960

DB 901 GCATGATAAACGAGAAAGCTCATAGACATATTAATTAATTAATTAATTAATTAATTAATTA 960

QY 961 TGAATAAATAATTTTATTTGAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAAAC 1020

DB 961 TGAATAAATAATTTTATTTGAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAAAC 1020

QY 1021 ACACCAATTTAAACCTTTTAAAAAGCATCTTAATAGAAACGAGGAAAGTTAAAGATTCAAC 1080

DB 1021 ACACCAATTTAAACCTTTTAAAAAGCATCTTAATAGAAACGAGGAAAGTTAAAGATTCAAC 1080

QY 1081 GAAGTGTTCGATAATGAAAAATGSGGTGGATTAGCAATTTGTTAAATGAATCAGGGTTAG 1140

DB 1081 GAAGTGTTCGATAATGAAAAATGSGGTGGATTAGCAATTTGTTAAATGAATCAGGGTTAG 1140

QY 1141 GATTAAATATTTAAATGAAAGAGGGAGAAATGAATGTTTAGAGTTTAAATGTGTCTTTTGG 1200

DB 1141 GATTAAATATTTAAATGAAAGAGGGAGAAATGAATGTTTAGAGTTTAAATGTGTCTTTTGG 1200

QY 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGGTCTTCCGTGGGCAGA 1260

DB 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGGTCTTCCGTGGGCAGA 1260

QY 1261 AGCGTCTTTTTTCGTTGGAAAAAATCTGCTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320

DB 1261 AGCGTCTTTTTTCGTTGGAAAAAATCTGCTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320

QY 1321 GTTGTACCGTCTCACAGAGAAAGAAAGCCCAACCCACCTTAAACCTCTCGATCC 1380

DB 1321 GTTGTACCGTCTCACAGAGAAAGAAAGCCCAACCCACCTTAAACCTCTCGATCC 1380

QY 1381 GACCGAGACTCTCTCCATTTTCCAGCGCGCACCGGAGAGCACGCGAGTTCGTTCGAC 1440

DB 1381 GACCGAGACTCTCTCCATTTTCCAGCGCGCACCGGAGAGCACGCGAGTTCGTTCGAC 1440

QY 1441 GAACAGGCTAGTGCAGTAGTTGTTGCTCGGAGATGCGCGGGAACCTGGAGAGAGCTGCC 1500

DB 1441 GAACAGGCTAGTGCAGTAGTTGTTGCTCGGAGATGCGCGGGAACCTGGAGAGAGCTGCC 1500

QY 1501 GTCGATGGAGCTCATACAGAGCTGCTCCGCGCATGAAGTGCAGCTCCAAGCCCGACAA 1560

DB 1501 GTCGATGGAGCTCATACAGAGCTGCTCCGCGCATGAAGTGCAGCTCCAAGCCCGACAA 1560

QY 1561 GCSCGTCTCTCGTTCGTTAAACCGCGCTCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1620

DB 1561 GCSCGTCTCTCGTTCGTTAAACCGCGCTCTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1620

QY 1621 TCGGTGTTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1680

DB 1621 TCGGTGTTTGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1680

QY 1681 GAGATGGGAAAGATTGCTGTGGCTTGTCTGGGATGTTGCGTTAGATCCGTGCAAAAGCTG 1740

DB 1681 GAGATGGGAAAGATTGCTGTGGCTTGTCTGGGATGTTGCGTTAGATCCGTGCAAAAGCTG 1740

QY 1741 TCTCCCTTTCTTTTGGTCAAAATCTGGGTTCTTCTTCCGCTCTCTCTCTCTCTCTCTCTCT 1800

DB 1741 TCTCCCTTTCTTTTGGTCAAAATCTGGGTTCTTCTTCCGCTCTCTCTCTCTCTCTCTCTCTCT 1800

QY 1801 TGCAAAACAAAAATGTTTGGCTTCGAGGGAAATCTTGTGTAATAATTTACTTGTGTTTCCACC 1860

Db 1801 TGCACACAAAAATGTTGGCTTCGAGGAAATCTTTGTAATAATTTACTTGTGTTCCACC 1860
Qy TTTGGAAACTGCGGCAATTTTGGATGGTCAAAATTTGCGTTTGTCTTGCACACTTGCATTT 1920
Db TTTGGAAACTGCGGCAATTTTGGATGGTCAAAATTTGCGTTTGTCTTGCACACTTGCATTT 1920
Qy GAGTCGATTTGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAAATCAAGCGAAAAAG 1980
Db GAGTCGATTTGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAAATCAAGCGAAAAAG 1980
Qy AAATATTATATGATCGGTGTCGACGCAATAATAGCAATTTCAAGCGTTTAAAGGTATAGA 2040
Db AAATATTATATGATCGGTGTCGACGCAATAATAGCAATTTCAAGCGTTTAAAGGTATAGA 2040
Qy ATTTATTTTCGTAATTCAGTAGTATGTTGAAATTTGGAGGTGCAATTTTGTCTCACACA 2100
Db ATTTATTTTCGTAATTCAGTAGTATGTTGAAATTTGGAGGTGCAATTTTGTCTCACACA 2100
Qy TGTATTATTTAAATTTGTCATGTCGCAATGGTGTGAGGATGAGCAAGAAACACAGGCG 2160
Db TGTATTATTTAAATTTGTCATGTCGCAATGGTGTGAGGATGAGCAAGAAACACAGGCG 2160
Qy GGTTCGTGGTTTTTGTCTCATGTGGAACCTTTTAAATAACCAAAACAGAAAGGTGCA 2220
Db GGTTCGTGGTTTTTGTCTCATGTGGAACCTTTTAAATAACCAAAACAGAAAGGTGCA 2220
Qy ACTTTATATGATGCAATAAGGTTTGGAAATGTCATCGGAATAAACAAGAGATAGA 2280
Db ACTTTATATGATGCAATAAGGTTTGGAAATGTCATCGGAATAAACAAGAGATAGA 2280
Qy CGTATTACTCTTGTGTTTCAATATTCACCTAAACCAAACTGATCTGTATGGTTATGTA 2340
Db CGTATTACTCTTGTGTTTCAATATTCACCTAAACCAAACTGATCTGTATGGTTATGTA 2340
Qy AACTGTGTGACACAAATTTCAAATTTCAAATTTGTTCTGTGCGCAATGCTCTGCTTCAG 2400
Db AACTGTGTGACACAAATTTCAAATTTCAAATTTGTTCTGTGCGCAATGCTCTGCTTCAG 2400
Qy CAATGCACTACTCTAAATTTTACTGGATCATCAAAACAAATCAATATGGAAGGTTTATTTCTA 2460
Db CAATGCACTACTCTAAATTTTACTGGATCATCAAAACAAATCAATATGGAAGGTTTATTTCTA 2460
Qy TTGTTCTTTTACGTTCAAGAATATAGCTAAATTTTAAACAACTCTCTGCACTACCTAATTT 2520
Db TTGTTCTTTTACGTTCAAGAATATAGCTAAATTTTAAACAACTCTCTGCACTACCTAATTT 2520
Qy TAGTGGACCGAATCTAGTCTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTTTCT 2580
Db TAGTGGACCGAATCTAGTCTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTTTCT 2580
Qy AATTCATGCAATGTCATGAGAACTGGTGTGTTTAAACAGCAATTTGTAAGAATGCCAGTAC 2640
Db AATTCATGCAATGTCATGAGAACTGGTGTGTTTAAACAGCAATTTGTAAGAATGCCAGTAC 2640
Qy ACCCTAATATGTTATTCAGAGATAGGATGACAGTGAACATAGAGCAAGTGTATGCA 2700
Db ACCCTAATATGTTATTCAGAGATAGGATGACAGTGAACATAGAGCAAGTGTATGCA 2700
Qy TTGATGCTTTTATAGTATGATCATCTGCCAGAAGCTTGAATATTTGATGTCATAGAATGG 2760
Db TTGATGCTTTTATAGTATGATCATCTGCCAGAAGCTTGAATATTTGATGTCATAGAATGG 2760
Qy TAAGCAATCTCTGGGTGTTCTAAACAAAAAAGGTAAAGATACAAATTTTGACATCCC 2820
Db TAAGCAATCTCTGGGTGTTCTAAACAAAAAAGGTAAAGATACAAATTTTGACATCCC 2820
Qy TTTATTTTACTAAATTTTAGTGTCCACTGCGTCCGAAAGGGAACACAGTCAACCGCTGAT 2880
Db TTTATTTTACTAAATTTTAGTGTCCACTGCGTCCGAAAGGGAACACAGTCAACCGCTGAT 2880
Qy TAAGATGAATTTTGTGTTGCCATTTAGCCACTGGTGATATCTTGAGGCTGCAAGTGGC 2940

Db 2881 TAAGATGAATTTTGTGTTGCCATTTAGCCACTGGTGATATGTTGAGGCTGCAAGTGGC 2940
Qy TGCTAAACCTCCACTTTGGGATTAAGGCTAAAGACTATGGACAAGGTAGTTTTTAAAGAA 3000
Db TGCTAAACCTCCACTTTGGGATTAAGGCTAAAGACTATGGACAAGGTAGTTTTTAAAGAA 3000
Qy ACATATAGCAACAGAAATATATAACAGAGGAATGGGTTTCTTGATCTTTTGTGTTCTTT 3060
Db ACATATAGCAACAGAAATATATAACAGAGGAATGGGTTTCTTGATCTTTTGTGTTCTTT 3060
Qy CTTTATCTTCTAGGAGAGCTTGTCTGATGACTTGGTTGTTGGGATTAATGATGAAGC 3120
Db CTTTATCTTCTAGGAGAGCTTGTCTGATGACTTGGTTGTTGGGATTAATGATGAAGC 3120
Qy CATGAAGAAACTTTCATGCCAGAAAGGTTTTATCTCTGATGGTTTTTCCCTAGAACCTGTTG 3180
Db CATGAAGAAACTTTCATGCCAGAAAGGTTTTATCTCTGATGGTTTTTCCCTAGAACCTGTTG 3180
Qy TCAAGCACAGAGGTGAGGTCTTGGTCAATATGCACCGCTATATAAAAGAGCTCTCTTTT 3240
Db TCAAGCACAGAGGTGAGGTCTTGGTCAATATGCACCGCTATATAAAAGAGCTCTCTTTT 3240
Qy TGTATTATAGAGTGTCTATATAAATGGAAGTCTTCTATCAITGTATCACTTTTCTTACTA 3300
Db TGTATTATAGAGTGTCTATATAAATGGAAGTCTTCTATCAITGTATCACTTTTCTTACTA 3300
Qy AAAATGTCAGCTGATGAAATGTTGGCCAAACAAAGTACTAAGATTGACAAGTGTCT 3360
Db AAAATGTCAGCTGATGAAATGTTGGCCAAACAAAGTACTAAGATTGACAAGTGTCT 3360
Qy AAAATTTGCAATTTGATGATCAATATCTGGAAGAAACGAATTTACCGGTGTTGGATCCACCC 3420
Db AAAATTTGCAATTTGATGATCAATATCTGGAAGAAACGAATTTACCGGTGTTGGATCCACCC 3420
Qy ATCAAGTGGTAGTCTTATCATACAAATTTGCTCTCTCTAGACTCTCTGAGCTTTGATGA 3480
Db ATCAAGTGGTAGTCTTATCATACAAATTTGCTCTCTCTAGACTCTCTGAGCTTTGATGA 3480
Qy TGTAAGTCAATACCAAGTACTCTGCTCTGCTTGCATTTCTCAGATCTCAGATTTTTTAC 3540
Db TGTAAGTCAATACCAAGTACTCTGCTCTGCTTGCATTTCTCAGATCTCAGATTTTTTAC 3540
Qy CATTTTTCATTTTCTTAATAGATTTGGTACATATAATGTTGATGCTTTGACAGCATATGC 3600
Db CATTTTTCATTTTCTTAATAGATTTGGTACATATAATGTTGATGCTTTGACAGCATATGC 3600
Qy GCCTTTTACATAATTTCCCTCTCTCATCGATCATCGTAGTACTCTGCTTACTTGT 3660
Db GCCTTTTACATAATTTCCCTCTCTCATCGATCATCGTAGTACTCTGCTTACTTGT 3660
Qy TTTAAGACAAACATGAGCCATCTTTTCAAAATTAAGGTTTCAATGAGTATG 3720
Db TTTAAGACAAACATGAGCCATCTTTTCAAAATTAAGGTTTCAATGAGTATG 3720
Qy TCCTCTGCCACTGTTTAAATAGTGTTCCTCCAAATAGTAAATAGTAGTGTGATAT 3780
Db TCCTCTGCCACTGTTTAAATAGTGTTCCTCCAAATAGTAAATAGTAGTGTGATAT 3780
Qy GGATATATTTTCTGTTTGGTTGTTCTGTCGACGAAGTAAATGCACTTTTGTG 3840
Db GGATATATTTTCTGTTTGGTTGTTCTGTCGACGAAGTAAATGCACTTTTGTG 3840
Qy ACAAGAAATGATAGGCAAGTGTTCGATGACACACTTTTCTCTGCAATATG 3900
Db ACAAGAAATGATAGGCAAGTGTTCGATGACACACTTTTCTCTGCAATATG 3900
Qy CCTGTTTAAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGATTTGATA 3960
Db CCTGTTTAAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGATTTGATA 3960
Qy TTGTTTCTTATTTGCTTCCAAATTAAGGATATCTCTTTAGCAATGATTTCAITGGGAC 4020
Db TTGTTTCTTATTTGCTTCCAAATTAAGGATATCTCTTTAGCAATGATTTCAITGGGAC 4020

QY 4021 ATATTGTGCTGCTATTAAAGTAAATTTGTTGATATCATATATATCTTTTAATTTGGTAAT 4080
DB |||||
QY 4021 ATATTGTGCTGCTATTAAAGTAAATTTGTTGATATCATATATATCTTTTAATTTGGTAAT 4080
DB |||||
QY 4081 ATTATGTGCACTTCTGCTCCCTGATTTGTTGCTTTTTCACAAAGGTTACTGAGAAACC 4140
DB |||||
QY 4081 ATTATGTGCACTTCTGCTCCCTGATTTGTTGCTTTTTCACAAAGGTTACTGAGAAACC 4140
DB |||||
QY 4141 TTAATTTCAAAGGAAGATGACACAGCTGCAGTATTGAAGTCAAGGCTTGAAGCTTCCAC 4200
DB |||||
QY 4141 TTAATTTCAAAGGAAGATGACACAGCTGCAGTATTGAAGTCAAGGCTTGAAGCTTCCAC 4200
DB |||||
QY 4201 GPACAACTAAGCCGTGATGTTTCTTTAGCAACTACGTTTTTAAATPATTCAGATATCT 4260
DB |||||
QY 4201 GPACAACTAAGCCGTGATGTTTCTTTAGCAACTAGCGTTTTTAAATPATTCAGATATCT 4260
DB |||||
QY 4261 TTTAGGATGTAGTCTGCTACTTCAGTTAAGGGCGGATTCCTTCAGTTGCATTACAGTGTTC 4320
DB |||||
QY 4261 TTTAGGATGTAGTCTGCTACTTCAGTTAAGGGCGGATTCCTTCAGTTGCATTACAGTGTTC 4320
DB |||||
QY 4321 TGTATATCTTTCATGTTTTT 4341
DB |||||
QY 4321 TGTATATCTTTCATGTTTTT 4341
DB |||||
RESULT 2
AAAA10960
ID AAA10960 standard; DNA; 4341 BP.
XX
AC AAA10960;
XX
DT 14-JUL-2000 (first entry)
XX
DE Rice adenylate kinase gene promoter fragment.
XX
KW Rice; adenylate kinase; promoter; gene expression control; antibacterial;
KW insecticide; light resistance; sucrose induction; flower promotion;
KW increase crop yield; herbicide resistance; ds.
XX
OS Oryza sativa.
XX
PN WO200015811-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1998; 98WO-JP04088.
XX
PR 10-SEP-1998; 98WO-JP04088.
XX
PA (NISC) NISSAN CHEM IND LTD.
XX
PI Uchimiya H, Arai S, Fushimi T, Tagawa M;
XX
DR WPI; 2000-271447/23.
XX
PT Rice adenylate kinase gene-originated DNA fragment with promoter
PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
PT induction in recombinant plants -
XX
PS Claim 1; Page 23-25; 30pp; Japanese.
XX
CC This sequence represents a rice adenylate kinase gene-originated DNA
CC fragment with promoter function in plants. The fragment can be used for
CC controlling the expression of desired foreign structural genes. The DNA
CC fragment is used to create a vector, which in turn is used to transform
CC bacteria or plant cells. The DNA fragment can be inserted into a plant to
CC control the expression of desired foreign structural genes, useful for
CC inducing in the plant e.g. herbicide, stress and light resistance,
CC production of antibacterial and insecticidal substances, sucrose
CC induction in recombinant plants, and in promoting flowering and crop
CC yield.

XX
SQ Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;
Query Match 100.0%; Score 4341; DB 21; Length 4341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCAGGAAGATTAAATTAGGTGGACACACAAACCCTGTGGTTGGTGGCCCTGTGTT 60
DB |||||
QY 1 CTGCAGGAAGATTAAATTAGGTGGACACACAAACCCTGTGGTTGGTGGCCCTGTGTT 60
DB |||||
QY 61 AATCAACTGGGTGTTGCTTGGACATGGTTTTGCAAGGAAATTAAGCAAGAAATTAAG 120
DB |||||
QY 61 AATCAACTGGGTGTTGCTTGGACATGGTTTTGCAAGGAAATTAAGCAAGAAATTAAG 120
DB |||||
QY 121 AAGAATCTCAAGCTGACATGAGAAACCGTAATCCAAATGGAAGCGAAATTTCAAGTCGTT 180
DB |||||
QY 121 AAGAATCTCAAGCTGACATGAGAAACCGTAATCCAAATGGAAGCGAAATTTCAAGTCGTT 180
DB |||||
QY 181 TCTTGTACTACCATGTTTGAATACATAGACAGTGCCCAACGTTTGAATGGCTCTATTGG 240
DB |||||
QY 181 TCTTGTACTACCATGTTTGAATACATAGACAGTGCCCAACGTTTGAATGGCTCTATTGG 240
DB |||||
QY 241 CTCGTGTGATCTGACTTGTGCACAAAGCATCAAAATTCCTTCTTGGAGTATCTTTATTA 300
DB |||||
QY 241 CTCGTGTGATCTGACTTGTGCACAAAGCATCAAAATTCCTTCTTGGAGTATCTTTATTA 300
DB |||||
QY 301 CCGAAACCCCAAGATATTCTTATTCACCTCAGGTTAAATTTGCTGNACTATGCAATG 360
DB |||||
QY 301 CCGAAACCCCAAGATATTCTTATTCACCTCAGGTTAAATTTGCTGNACTATGCAATG 360
DB |||||
QY 361 AATACAAATTCGCAAAATATCATGTTTATCTATCTTGTCTCAAAATTTGAGTCCAA 420
DB |||||
QY 361 AATACAAATTCGCAAAATATCATGTTTATCTATCTTGTCTCAAAATTTGAGTCCAA 420
DB |||||
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTTAATTTTTCATGAAA 480
DB |||||
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTTAATTTTTCATGAAA 480
DB |||||
QY 481 CGCAATTCACACCGTTCGAGAAATATCTGTCATAATAGTAGTCTAGTCGAGAAACAAA 540
DB |||||
QY 481 CGCAATTCACACCGTTCGAGAAATATCTGTCATAATAGTAGTCTAGTCGAGAAACAAA 540
DB |||||
QY 541 ATTAATATCATATAAAAGAGGTTGTTAATACAAACCATGTTTCGTACTACAACTCT 600
DB |||||
QY 541 ATTAATATCATATAAAAGAGGTTGTTAATACAAACCATGTTTCGTACTACAACTCT 600
DB |||||
QY 601 AATTTGTAAATTTCTTATTTTCAGTCACAAATTTCCAATTTTCCAAATTAAGAAATTAACGT 660
DB |||||
QY 601 AATTTGTAAATTTCTTATTTTCAGTCACAAATTTCCAATTTTCCAAATTAAGAAATTAACGT 660
DB |||||
QY 661 AGACGGCTAAGCCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGACGAAAAATA 720
DB |||||
QY 661 AGACGGCTAAGCCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGACGAAAAATA 720
DB |||||
QY 721 TGATGGTTTATTAATATGATTTTAAATTAACCTTTACATAAAATTTCTTTAGGAAACA 780
DB |||||
QY 721 TGATGGTTTATTAATATGATTTTAAATTAACCTTTACATAAAATTTCTTTAGGAAACA 780
DB |||||
QY 781 TATCATTTAATGTTTGAAGAACGTCACATAGAAACCTAAGAACGATAGTTCGGAAA 840
DB |||||
QY 781 TATCATTTAATGTTTGAAGAACGTCACATAGAAACCTAAGAACGATAGTTCGGAAA 840
DB |||||
QY 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAAC 900
DB |||||
QY 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAAC 900
DB |||||
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTATTACTTAGATATTATAAATAACT 960
DB |||||
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTATTACTTAGATATTATAAATAACT 960
DB |||||
QY 961 TGAAGAAAAATTTATTGAAATTTTAAACAATGTATGCATAAAATTTTAAAAAAC 1020
DB |||||

Db 961 TGA AAAAATAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATATTTTAAAAAC 1020
Qy 1021 ACACCAATTTAACCCTTTAAAGCATCTTAATAGGAAACGAGAAAGTTAAAGATTCCACC 1080
Db 1021 ACACCAATTTAACCCTTTAAAGCATCTTAATAGGAAACGAGAAAGTTAAAGATTCCACC 1080
Qy 1081 GAACTGTTTGGATAATGAAATAATGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
Db 1081 GAACTGTTTGGATAATGAAATAATGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
Qy 1141 GATTAAATATTAATAAGAGAGAGAGAAATGAATGGTTAGAGTTTAAATGTGTCTTTTGG 1200
Db 1141 GATTAAATATTAATAAGAGAGAGAGAAATGAATGGTTAGAGTTTAAATGTGTCTTTTGG 1200
Qy 1201 GTGGGTAGAAAATATTTTGCATACACTCCCGAGAGCGGTGGCTTGCCTGGGCGAGA 1260
Db 1201 GTGGGTAGAAAATATTTTGCATACACTCCCGAGAGCGGTGGCTTGCCTGGGCGAGA 1260
Qy 1261 AGCGTCTTTTTCGTTGGAAAAAAAACCTGCTTAAAAAGGAAAAAGAGAGAGAGAGAGAGAG 1320
Db 1261 AGCGTCTTTTTCGTTGGAAAAAAAACCTGCTTAAAAAGGAAAAAGAGAGAGAGAGAGAG 1320
Qy 1321 GTTGTACCGTCTCACAGAGAAACGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GTTGTACCGTCTCACAG 1380
Qy 1381 GACCGAGACTCTCCATTTTCAAGCGGCGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 GACCGAGACTCTCCATTTTCAAGCGGCGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 GAACAAGGCTAGTGCAGTAGTTGTGTCGGAGATGGCGGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GAACAAGGCTAGTGCAGTAGTTGTGTCGGAGATGGCGGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1501 GTCCATGAGCTGATGACGAGAGCTGCTCCGCGCATGAAAGTGCAGCTCCAAAGAGAGAGAG 1560
Db 1501 GTCCATGAGCTGATGACGAGAGCTGCTCCGCGCATGAAAGTGCAGCTCCAAAGAGAGAGAG 1560
Qy 1561 GCGCGTCATCTCTCGTGGTAAACCGCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTC 1620
Db 1561 GCGCGTCATCTCTCGTGGTAAACCGCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTC 1620
Qy 1621 TCAGTGGTTTGGTTGGTTGGTGGATCCGGTTCCGGGCGCGCTGCTGCGCTGGTGGG 1680
Db 1621 TCAGTGGTTTGGTTGGTTGGTGGATCCGGTTCCGGGCGCGCTGCTGCGCTGGTGGG 1680
Qy 1681 GAGATGGGAAAGATTCCTGTGGCTTGTGGGATGTTGGTTAGATCCGTGCAAGAGCTG 1740
Db 1681 GAGATGGGAAAGATTCCTGTGGCTTGTGGGATGTTGGTTAGATCCGTGCAAGAGCTG 1740
Qy 1741 TCTCCCTTTCTTTTGGGTCAAAATCTGGGTTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1800
Db 1741 TCTCCCTTTCTTTTGGGTCAAAATCTGGGTTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1800
Qy 1801 TGCAAAACAAATTTTGGCTTCGAGGAAATCTTGTAAATTTACTTGTGTGTTCACCC 1860
Db 1801 TGCAAAACAAATTTTGGCTTCGAGGAAATCTTGTAAATTTACTTGTGTGTTCACCC 1860
Qy 1861 TTTGGAACCTCGCGCAATTTTGGTCAAAATTTGGTGGTGTCTTGTGCAACTTGGCAT 1920
Db 1861 TTTGGAACCTCGCGCAATTTTGGTCAAAATTTGGTGGTGTCTTGTGCAACTTGGCAT 1920
Qy 1921 GAGTCGATTTGGGTTTTCATTTGACGAGAGATATAGTGACAGAAATCAAGCGGAAAG 1980
Db 1921 GAGTCGATTTGGGTTTTCATTTGACGAGAGATATAGTGACAGAAATCAAGCGGAAAG 1980
Qy 1981 AAATATTTATATGATCGGTGTCGAGCATAATAGGAAATTCAGGCGTTAAAGGTATAGA 2040
Db 1981 AAATATTTATATGATCGGTGTCGAGCATAATAGGAAATTCAGGCGTTAAAGGTATAGA 2040
Qy 2041 ATTTATTCGATTCAGTAGTATGTGTTGAAATTTGGAGGTTGCAATTTTGTCTCTCACACA 2100
Db 2041 ATTTATTCGATTCAGTAGTATGTGTTGAAATTTGGAGGTTGCAATTTTGTCTCTCACACA 2100

Qy 2101 TGTATATTTTAAATTTGTGCATGTGCAATGGTGTAGGATGAGCAAGAAACAAACAGGCG 2160
Db 2101 TGTATATTTTAAATTTGTGCATGTGCAATGGTGTAGGATGAGCAAGAAACAAACAGGCG 2160
Qy 2161 GGTTCGTTGGTTTGTCTGATGTGAAACTTTTAAATAAACCAAAACAGAAAGGTGCA 2220
Db 2161 GGTTCGTTGGTTTGTCTGATGTGAAACTTTTAAATAAACCAAAACAGAAAGGTGCA 2220
Qy 2221 ACTTTATATGTATCGATAAAGTTTGGAAATGTCACCTCGGAATAAAAAACAAAGAGATAGA 2280
Db 2221 ACTTTATATGTATCGATAAAGTTTGGAAATGTCACCTCGGAATAAAAAACAAAGAGATAGA 2280
Qy 2281 CGTATTAATCTCTTGTTCATTAATTCACACTAAACCAAAATCGATCTGTATGTTATGTA 2340
Db 2281 CGTATTAATCTCTTGTTCATTAATTCACACTAAACCAAAATCGATCTGTATGTTATGTA 2340
Qy 2341 AACTGTGTGACAAACAAATTTCAAATAAACTTTCTGTGCGCAATGCTCTGCTTCAG 2400
Db 2341 AACTGTGTGACAAACAAATTTCAAATAAACTTTCTGTGCGCAATGCTCTGCTTCAG 2400
Qy 2401 CAATGCACTACTCTAAATTTTACTGGAATCATCAAAACAAATCATATGGAAGGTTTATTTCTA 2460
Db 2401 CAATGCACTACTCTAAATTTTACTGGAATCATCAAAACAAATCATATGGAAGGTTTATTTCTA 2460
Qy 2461 TTGCTTCTTTTACGTTTCATGAAGTATAGCTTAATTTACAAACTCTCTGCACTACCTAATTT 2520
Db 2461 TTGCTTCTTTTACGTTTCATGAAGTATAGCTTAATTTACAAACTCTCTGCACTACCTAATTT 2520
Qy 2521 TAGTGGACCGAATCTAGTCACTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTGTCT 2580
Db 2521 TAGTGGACCGAATCTAGTCACTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTGTCT 2580
Qy 2581 AATTCATGCACTGTCACCTGAGAACCTGGTGTGTTAACAGCATTTGAAGATGCCAGTAC 2640
Db 2581 AATTCATGCACTGTCACCTGAGAACCTGGTGTGTTAACAGCATTTGAAGATGCCAGTAC 2640
Qy 2641 ACCCTAATATGTTATTGAGGATAGGATAGTGCAGTGTAAACATAGAGCAAGTGTATGCAA 2700
Db 2641 ACCCTAATATGTTATTGAGGATAGGATAGTGCAGTGTAAACATAGAGCAAGTGTATGCAA 2700
Qy 2701 TTGATGCTTTAGATAGTTTCAATAGCATCTGCGAGAACCTTGAATATATGATGCTAGAAAG 2760
Db 2701 TTGATGCTTTAGATAGTTTCAATAGCATCTGCGAGAACCTTGAATATATGATGCTAGAAAG 2760
Qy 2761 TAAGACAACTCTGCGGTGTTCTAAAAAAGGTAAAGACATACAAATTTTGACATCCC 2820
Db 2761 TAAGACAACTCTGCGGTGTTCTAAAAAAGGTAAAGACATACAAATTTTGACATCCC 2820
Qy 2821 TTTATTTTACTAAATTTTAGGTCACCTGGCTCGGAAAGGGAACACAGTCCCGCTGAT 2880
Db 2821 TTTATTTTACTAAATTTTAGGTCACCTGGCTCGGAAAGGGAACACAGTCCCGCTGAT 2880
Qy 2881 TAAGGATGAATTTGCTGTGCGCATTTAGCCATGAGTATGTTGAGGCTGCGATGGC 2940
Db 2881 TAAGGATGAATTTGCTGTGCGCATTTAGCCATGAGTATGTTGAGGCTGCGATGGC 2940
Qy 2941 TGCTAAACCTCCACTTGGGATTAAGGCTAAAGGCTATGGAAGGAGTATTTTAAAGAA 3000
Db 2941 TGCTAAACCTCCACTTGGGATTAAGGCTAAAGGCTATGGAAGGAGTATTTTAAAGAA 3000
Qy 3001 ACATATAGCAACAGAAATTTAACCAGAGAGGATGGGTTTCTTGATCTTTTGTCTTTT 3060
Db 3001 ACATATAGCAACAGAAATTTAACCAGAGAGGATGGGTTTCTTGATCTTTTGTCTTTT 3060
Qy 3061 CCTTATCTTCTAGGAGAGCTTGTTCATGATGCTTGTGTTGGATTTTATGATGAAGC 3120
Db 3061 CCTTATCTTCTAGGAGAGCTTGTTCATGATGCTTGTGTTGGATTTTATGATGAAGC 3120
Qy 3121 CATGAAGAAACCTCATGCCAGAAAGGTTTATCTTGTATGGTTCCCTAGAACTGTTGT 3180
Db 3121 CATGAAGAAACCTCATGCCAGAAAGGTTTATCTTGTATGGTTCCCTAGAACTGTTGT 3180

```
QY 3181 TCAAGCACAGAAGGTGAGTCTTGGTCAATATGACCGCTATATAAAGAGCTCCTTTT 3240
Db 3181 TCAAGCACAGAAGGTGAGTCTTGGTCAATATGACCGCTATATAAAGAGCTCCTTTT 3240
QY 3241 TGTATTAGAGCTGTCTATATAAATGAAGAGTTTCTTATCATTTGTATCACTTTCTTACTA 3300
Db 3241 TGTATTAGAGCTGTCTATATAAATGAAGAGTTTCTTATCATTTGTATCACTTTCTTACTA 3300
QY 3301 AAAATGGTGCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCT 3360
Db 3301 AAAATGGTGCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCT 3360
QY 3361 AAATTTGGCAATTGATGATCAATACTGGAAGAACGAAATACCGGTCGTTGGATCCACCC 3420
Db 3361 AAATTTGGCAATTGATGATCAATACTGGAAGAACGAAATACCGGTCGTTGGATCCACCC 3420
QY 3421 ATCAAGTGTAGATCTTATCATACAAAATTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3480
Db 3421 ATCAAGTGTAGATCTTATCATACAAAATTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3480
QY 3481 TGTAAAGTCATACCAGATTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3540
Db 3481 TGTAAAGTCATACCAGATTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3540
QY 3541 CATTTTTCATTATTTCTAATAGATTGGTACATATAAATGGTTGATGCTTGCAGCATATGC 3600
Db 3541 CATTTTTCATTATTTCTAATAGATTGGTACATATAAATGGTTGATGCTTGCAGCATATGC 3600
QY 3601 GCCTTTACCATAATTTCCCTGCTCATCATCGATCATCTGATCTCTCTCTCTCTCTCTCTCT 3660
Db 3601 GCCTTTACCATAATTTCCCTGCTCATCATCGATCATCTGATCTCTCTCTCTCTCTCTCTCT 3660
QY 3661 TTTAAGAACAAACATGAGCCATCACTTTCTCAAAAATAAAAAGGTTCAITTGAGTACCT 3720
Db 3661 TTTAAGAACAAACATGAGCCATCACTTTCTCAAAAATAAAAAGGTTCAITTGAGTACCT 3720
QY 3721 TCCTCTTGCACCTGTTAATAGTTAGTCTTCCCAATAGTTAAATAGGTAGTGTGATAT 3780
Db 3721 TCCTCTTGCACCTGTTAATAGTTAGTCTTCCCAATAGTTAAATAGGTAGTGTGATAT 3780
QY 3781 GGATATATTTTCTTGTGTTGTTCTCGTACGAAGTAAATGACCTTTTGTG 3840
Db 3781 GGATATATTTTCTTGTGTTGTTCTCGTACGAAGTAAATGACCTTTTGTG 3840
QY 3841 ACAAGAAATGATAGGAGTGTGATGACACACATTTTCTGTTCTGACAAATTATG 3900
Db 3841 ACAAGAAATGATAGGAGTGTGATGACACACATTTTCTGTTCTGACAAATTATG 3900
QY 3901 CCTGTTTAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGAGATTGTATA 3960
Db 3901 CCTGTTTAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGAGATTGTATA 3960
QY 3961 TTGTTTCTATTCTTCCCAATTAAGCATATTTCTTTAGCAATGATTTTCATGTGGAC 4020
Db 3961 TTGTTTCTATTCTTCCCAATTAAGCATATTTCTTTAGCAATGATTTTCATGTGGAC 4020
QY 4021 ATATTTGCTGCTATTAAGTAAATTTGTTGATATCATATATCTTTTAATTTGGTAAAT 4080
Db 4021 ATATTTGCTGCTATTAAGTAAATTTGTTGATATCATATATCTTTTAATTTGGTAAAT 4080
QY 4081 ATTATGTGCATCTTCTCTCTCTGATTTGTTCTTTTTCACAAAGGTTACTCGAAGACCC 4140
Db 4081 ATTATGTGCATCTTCTCTCTCTGATTTGTTCTTTTTCACAAAGGTTACTCGAAGACCC 4140
QY 4141 TTAATTCAAAGAAAGATGACACAGCTCGAGTATTGAAGTCAAGGCTTGAAGCCTTCCAC 4200
Db 4141 TTAATTCAAAGAAAGATGACACAGCTCGAGTATTGAAGTCAAGGCTTGAAGCCTTCCAC 4200
QY 4201 GTACAACTAAGCCTGTATGTTTCTTTTAGCAACTACGTTTTTAATATTTCAGATATTCT 4260
Db 4201 GTACAACTAAGCCTGTATGTTTCTTTTAGCAACTACGTTTTTAATATTTCAGATATTCT 4260
QY 4261 TTTAGGATGTAGTCGTACTTTCAGTTAAGCGGATTCCTTTCAGTTGCATTACAGTGTCTCC 4320
```

```
Db 4261 TTTAGGATGTAGTCGTACTTTCAGTTAAGCGGATTCCTTTCAGTTGCATTACAGTGTCTCC 4320
QY 4321 TGTATATCTTTTCATTTGTTTTT 4341
Db 4321 TGTATATCTTTTCATTTGTTTTT 4341

RESULT 3
ID AAA07482 standard; DNA; 1440 BP.
XX AAA07482;
XX 17-JUL-2000 (first entry)
XX Rice adenylate kinase DNA sequence fragment #2.
XX Rice; adenylate kinase; promoter; structural gene expression;
KW plant gene expression; herbicide resistance; stress resistance;
KW light resistance; antibacterial production; flowering induction;
KW insecticidal substance production; ss.
XX Oryza sativa.
XX WO200015812-A1.
XX 23-MAR-2000.
XX 08-SEP-1999; 99WO-JP04847.
XX 10-SEP-1998; 98WO-JP04088.
XX (NISC) NISSAN CHEM IND LTD.
XX Uchimiya H, Arai S, Fushimi T, Tagawa M, Fukuzawa H;
XX WPI; 2000-271448/23.
XX Rice adenylate kinase gene-originated DNA fragment with promoter
PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
PT induction in recombinant plants -
XX Claim 10; Page 28; 33pp; Japanese.
XX This sequence represents a fragment of the rice adenylate kinase gene.
CC The invention relates to a DNA fragment (such as this sequence) that has
CC a regulatory function (preferably acting as a promoter) on the expression
CC of desired foreign structural genes in a plant when introduced into a
CC plant body or plant cell. The DNA fragment can be inserted into a plant
CC to control the expression of desired foreign structural genes, useful for
CC inducing in the plant e.g. herbicide, stress and light resistance,
CC production of antibacterial and insecticidal substances, sucrose,
CC induction in recombinant plants, and in promoting flowering and crop
CC yield. The DNA fragment has a regulatory function on the expression of
CC desired structural genes in a plant, particularly with control of the
CC expression in specific tissues and locations to provide safe chemicals.
XX Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

Query Match 33.2%; Score 1440; DB 21; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACGGAAGATTAAATAGGTGGACACACCAACCCTGTGGTTGGTGACGCCCTGTGTT 60
Db 1 CTGACGGAAGATTAAATAGGTGGACACACCAACCCTGTGGTTGGTGACGCCCTGTGTT 60
QY 61 AATCAACTGGGGTGTCTCGACATGTTTTTTCAGGAAATTAAGCAAGAAATTAAG 120
Db 61 AATCAACTGGGGTGTCTCGACATGTTTTTTCAGGAAATTAAGCAAGAAATTAAG 120
```


QY 121 AAGATGCTCAAGCTGACATGAGAAAACGTAATCAATCGAAGCGAATTTCAAGTCGTTTC 180
Db 121 AAGATGCTCAAGCTGACATGAGAAAACGTAATCAATCGAAGCGAATTTCAAGTCGTTTC 180
QY 181 TCTTGTAAGTACTGATGTTTGAATACATAGACAGTGCACCAAGTTCGATGGCTCCTATTGG 240
Db 181 TCTTGTAAGTACTGATGTTTGAATACATAGACAGTGCACCAAGTTCGATGGCTCCTATTGG 240
QY 241 CTGCTGTGATGATGCTGTTGTGCACAAAGCATCAAAATGCTTCCTGGAGTATCTTTATTA 300
Db 241 CTGCTGTGATGATGCTGTTGTGCACAAAGCATCAAAATGCTTCCTGGAGTATCTTTATTA 300
QY 301 CCGAAAACCCCAAGATTTATCTTATTCACCTCAGGGTAAATGCTGCAACTATGCAATG 360
Db 301 CCGAAAACCCCAAGATTTATCTTATTCACCTCAGGGTAAATGCTGCAACTATGCAATG 360
QY 361 AATACAAATTCGAAATATATGTTGTTATCTATCTTCTCAAAATGAAATTTGAGTCCAA 420
Db 361 AATACAAATTCGAAATATATGTTGTTATCTATCTTCTCAAAATGAAATTTGAGTCCAA 420
QY 421 CTGAGCTGCAATACGATTTCTTTTCAAAAAGAAATTTAAATTTTCTTATGAAA 480
Db 421 CTGAGCTGCAATACGATTTCTTTTCAAAAAGAAATTTAAATTTTCTTATGAAA 480
QY 481 CGCAATTCACACCGTTTCGAGAAATATGCTGTCATAAATAAGTAGTCTAGTGCAGAAACAA 540
Db 481 CGCAATTCACACCGTTTCGAGAAATATGCTGTCATAAATAAGTAGTCTAGTGCAGAAACAA 540
QY 541 ATTAATATCATATAAAGAGAGTTGTTAAATTAACAACATGTTTCGTAACAACTCT 600
Db 541 ATTAATATCATATAAAGAGAGTTGTTAAATTAACAACATGTTTCGTAACAACTCT 600
QY 601 AATTTGTAATTTCTATTTTCAGTCACAAAATTCCTCAATTTCCAAATTAAGAAAATAAAGCT 660
Db 601 AATTTGTAATTTCTATTTTCAGTCACAAAATTCCTCAATTTCCAAATTAAGAAAATAAAGCT 660
QY 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAATA 720
Db 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAATA 720
QY 721 TGATGGTTTAAATTAATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
Db 721 TGATGGTTTAAATTAATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
QY 781 TATCATTTAATGTTTGAAGAAAGCTGCACATAAGAAAACCTAAGAACGATGAGTTGGAAA 840
Db 781 TATCATTTAATGTTTGAAGAAAGCTGCACATAAGAAAACCTAAGAACGATGAGTTGGAAA 840
QY 841 CAAGAGAAAACACAGCCTTAAGGCTTCTGATCCTCTAGTTGGAGTTGATTTTCAAAC 900
Db 841 CAAGAGAAAACACAGCCTTAAAGGCTTCTGATCCTCTAGTTGGAGTTGATTTTCAAAC 900
QY 901 GCATGATTAACGAGAAAGCTCATTAGCACATTAATTAATTAATTAATTAATTAATTAATTA 960
Db 901 GCATGATTAACGAGAAAGCTCATTAGCACATTAATTAATTAATTAATTAATTAATTAATTA 960
QY 961 TGAAGAAAATATTTATTTGAATTTTAAACAATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TGAAGAAAATATTTATTTGAATTTTAAACAATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 ACACCAATTTAACCTTTAAAGAGCATCTTAATAGGAACGAGGAGTTAAGATTCACC 1080
Db 1021 ACACCAATTTAACCTTTAAAGAGCATCTTAATAGGAACGAGGAGTTAAGATTCACC 1080
QY 1081 GAAGTGTGTAATGAAATGAGGAGGATTAAGATTTGATGATGATGATGATGATGATGATGATG 1140
Db 1081 GAAGTGTGTAATGAAATGAGGAGGATTAAGATTTGATGATGATGATGATGATGATGATGATG 1140
QY 1141 GATTAAATATTAATGAAAGAGGAGGATTAAGATTTGATGATGATGATGATGATGATGATGATG 1200
Db 1141 GATTAAATATTAATGAAAGAGGAGGATTAAGATTTGATGATGATGATGATGATGATGATGATG 1200
QY 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGGAGAGCGGTGCTGCTGCGTGGGCGAGA 1260

Db 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGGAGAGCGGTGCTGCTGCGTGGGCGAGA 1260
QY 1261 AGGCTCTTTTTCGTTGGAAAAAACTGCTTTAAAAAGGAAAAACAGAAAGAGCCAGCTTTT 1320
Db 1261 AGGCTCTTTTTCGTTGGAAAAAACTGCTTTAAAAAGGAAAAACAGAAAGAGCCAGCTTTT 1320
QY 1321 GTTGTCACTGCTCTCACAGAAAGGAAACAAAGGCGCCACACCTTAAACCTCTCGATCC 1380
Db 1321 GTTGTCACTGCTCTCACAGAAAGGAAACAAAGGCGCCACACCTTAAACCTCTCGATCC 1380
QY 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGGGAGAGCACGCGAGTTCGTCGAC 1440
Db 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGGGAGAGCACGCGAGTTCGTCGAC 1440
RESULT 4
AAA10961
ID AAA10961 standard; DNA; 1440 BP.
XX
AC AAA10961;
XX
DT 14-JUL-2000 (first entry)
XX
DE Partial rice adenylate kinase gene promoter fragment.
KW Rice; adenylate kinase; promoter; gene expression control; antibacterial;
KW insecticide; light resistance; sucrose induction; flower promotion;
KW increase crop yield; herbicide resistance; ds.
XX
OS Oryza sativa.
XX
PN WO200015811-A1.
XX
PD 23-MAR-2000.
XX
PF 10-SEP-1998; 98WO-JP04088.
XX
PR 10-SEP-1998; 98WO-JP04088.
XX
PA (NISC) NISSAN CHEM IND LTD.
XX
PI Uchimiya H, Arai S, Fushimi T, Tagawa M;
XX
DR WPI; 2000-271447/23.
XX
PT Rice adenylate kinase gene-originated DNA fragment with promoter
PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
PT induction in recombinant plants -
XX
PS Example 4; Page 26; 30pp; Japanese.
XX
CC This sequence represents a partial rice adenylate kinase gene-originated
CC DNA fragment with promoter function in plants. The fragment can be used
CC for controlling the expression of desired foreign structural genes. The
CC DNA fragment is used to create a vector, which in turn is used to
CC transform bacteria or plant cells. The DNA fragment can be inserted into
CC a plant to control the expression of desired foreign structural genes,
CC useful for inducing in the plant e.g. herbicide, stress and light
CC resistance, production of antibacterial and insecticidal substances,
CC sucrose induction in recombinant plants, and in promoting flowering and
CC crop yield.
XX
SQ Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

Query Match 33.2%; Score 1440; DB 21; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAGGAAGATTAATTAAGTGGACACCAAAACCTGTGGTGGACCCCTGTGTT 60
Db 1 CTGAGGAAGATTAATTAAGTGGACACCAAAACCTGTGGTGGTGGACCCCTGTGTT 60

```
QY 61 AATCAACTGGGGTTCCTTGGACATGGTTTTGACGAAAAATTAAGCAAGAAAAATTAAG 120
Db 61 AATCAACTGGGGTTCCTTGGACATGGTTTTGACGAAAAATTAAGCAAGAAAAATTAAG 120
QY 121 AGAATGCTCAAGCTGACATGAGAAAACGTAATCCAAATGAAGCAATTTCAAGTCGTTTC 180
Db 121 AGAATGCTCAAGCTGACATGAGAAAACGTAATCCAAATGAAGCAATTTCAAGTCGTTTC 180
QY 181 TCTTGTAATACCATGTTTGAATACATACACAGATGCCACGTTTGATGGCTCCTATTGG 240
Db 181 TCTTGTAATACCATGTTTGAATACATACACAGATGCCACGTTTGATGGCTCCTATTGG 240
QY 241 CTCGTGTGATPACTGACTTGTGTGCACAAAGCATCAAAATGCTTCTTGAGTATCTTTATTA 300
Db 241 CTCGTGTGATPACTGACTTGTGTGCACAAAGCATCAAAATGCTTCTTGAGTATCTTTATTA 300
QY 301 CCGAAAAACCCCAAGATTAATCTATCCACCTCAGGGTAATTTGTGCTGAACATGCAATG 360
Db 301 CCGAAAAACCCCAAGATTAATCTATCCACCTCAGGGTAATTTGTGCTGAACATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGGTTTATCTATCTGCTCAAAATGAAATTCAGTCCAA 420
Db 361 AATACAAATTCGCAAAATATCATGGTTTATCTATCTGCTCAAAATGAAATTCAGTCCAA 420
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAAGAAATTAATTAATTTTTTTTTTCATGAAA 480
Db 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAAGAAATTAATTAATTTTTTTTTTCATGAAA 480
QY 481 CGCAATTCACACCGTTCGAGAAATATGCTGCATATAATAGTAGTCTAGTGCAGAAACAAA 540
Db 481 CGCAATTCACACCGTTCGAGAAATATGCTGCATATAATAGTAGTCTAGTGCAGAAACAAA 540
QY 541 ATTAATATACATAAAAAAGAGGTTGTTAATTAACAACCATGTTTCGTACTACAACTCT 600
Db 541 ATTAATATACATAAAAAAGAGGTTGTTAATTAACAACCATGTTTCGTACTACAACTCT 600
QY 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAAATTTCCAAATTAAGAAAAATAACGT 660
Db 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAAATTTCCAAATTAAGAAAAATAACGT 660
QY 661 AGACGGCTAAGCCACCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
Db 661 AGACGGCTAAGCCACCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
QY 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTTCAATATAATTTCTTTAGGAAACA 780
Db 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTTCAATATAATTTCTTTAGGAAACA 780
QY 781 TATCATTTAATGGTTTGAANAACGTCACATAGNAACCTAGAACGATGAGTTGGGAAA 840
Db 781 TATCATTTAATGGTTTGAANAACGTCACATAGNAACCTAGAACGATGAGTTGGGAAA 840
QY 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCCCTCTAGTTGGAGGTTGATTTTCAAAAC 900
Db 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCCCTCTAGTTGGAGGTTGATTTTCAAAAC 900
QY 901 GCATGATAACGAGAAAGCTCATTACACATATTAATCTAGATATTTAATTAATAAAT 960
Db 901 GCATGATAACGAGAAAGCTCATTACACATATTAATCTAGATATTTAATTAATAAAT 960
QY 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGATGATCAATAATTTTAAAAAAC 1020
Db 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGATGATGATCAATAATTTTAAAAAAC 1020
QY 1021 ACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGAAAACGAGGAAGTTAAAGATTCACC 1080
Db 1021 ACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGAAAACGAGGAAGTTAAAGATTCACC 1080
QY 1081 GAAAGTGTGGGATAATGAAAAATGGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
Db 1081 GAAAGTGTGGGATAATGAAAAATGGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
```

```
QY 1141 GATTAATAATTTAAAAATGAAGAGGAGAGAAATGAATGTTAGAGTTTAAATGTCGCTTTTGG 1200
Db 1141 GATTAATAATTTAAAAATGAAGAGGAGAGAAATGAATGTTAGAGTTTAAATGTCGCTTTTGG 1200
QY 1201 GTGGGTAGAAAAATTAATTTGCCATACACTCCCGAGAGCGGTGCGTCTTGGTGGGCAGA 1260
Db 1201 GTGGGTAGAAAAATTAATTTGCCATACACTCCCGAGAGCGGTGCGTCTTGGTGGGCAGA 1260
QY 1261 AGCGTCTTTTTCGTTGGAAAAAAAACCTCTTAAAAAGGAAAAACAGAAAGAGCCAGCTTT 1320
Db 1261 AGCGTCTTTTTCGTTGGAAAAAAAACCTCTTAAAAAGGAAAAACAGAAAGAGCCAGCTTT 1320
QY 1321 GTTGTACCGCTCTCACAGAAAACGAAACAAAAAGCCCAACCTTAAACCTCTCTCGATCC 1380
Db 1321 GTTGTACCGCTCTCACAGAAAACGAAACAAAAAGCCCAACCTTAAACCTCTCTCGATCC 1380
QY 1381 GACCGAGACTCTCCATTTTCAGGGGGGACGCGGAGAGACGCGGAGGAGTTCTGTCGAC 1440
Db 1381 GACCGAGACTCTCCATTTTCAGGGGGGACGCGGAGAGACGCGGAGGAGTTCTGTCGAC 1440

RESULT 5
AAC48619
ID AAC48619 standard; DNA; 809 BP.
XX
AC AAC48619;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58135.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; terminator; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
```

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140581.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01455913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 03-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

```
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 2.7%; Score 115.8; DB 21; Length 809;
Matches 141; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 3308 GTCGAGCTTCATGAAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTT 3367
Db 421 GAGAAGCTCGATGAGATGCTTAAGAGCGGAGGAATCAAAATTGACAAAGTTCTCAACTTT 480
QY 3368 GCAATTGATGATCAATCACTACTGGAAGAACGAATTACCGGTGCTTGGATCCACCCATCAAGT 3427
Db 481 GCTATTGATGACCAATCTTTGGAGGAAGAATAACCGGGGATGGATCCACCCATCGAGT 540
QY 3428 GGTAGATCTTATCATACAAATTTGCTCTCCCTTAAGACTCCTGGACTTGATGATGAAGT 3487
Db 541 GGCAGGAGTTACCAACACCAAAATTGCTCTCTCCCAAAACCCCTGGAGTTGATGATGATT 600
QY 3488 CAT 3490
Db 601 GAT 603

RESULT 6
AAC48110
ID AAC48110 standard; DNA; 482 BP.
AC AAC48110;
XX
DT 18-OCT-2000 (first entry)
DE
XX
DE Zea mays DNA fragment SEQ ID NO: 56299.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
```

| | | |
|----|--------------|---------------|
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161406. |
| PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161920. |
| PR | 28-OCT-1999; | 99US-0161922. |
| PR | 28-OCT-1999; | 99US-0161993. |
| PR | 23-OCT-1999; | 99US-0162142. |

Query Match 2.7%; Score 115.2; DB 21; Length 482;
 Best Local Similarity 80.4%; Pred. No. 3e-16;
 Matches 135; Conservative 0; Mismatches 33; Indels 0; Gaps

| | | | |
|----|------|--|------|
| QY | 2838 | TAGTCCACTTGGCTCGGAAAGGAAACACAGTCACCGCTGATTAAAGGATGAATTTTGCT | 2897 |
| Db | 271 | TCGGCCACCTGGCTCTGGAAGGGTACTCAGTCTCCCTTATTAAAGGATGAATATTGCC | 330 |
| QY | 2898 | TGTGCCATTTAGCCACTTGGTGATATGTTGAGGGCTGCAGTGGCTGTAAAACTCCACTTG | 2957 |
| Db | 331 | TGTGCCATTTAGCCACTTGGTGATATGTTGAGGGCTGTGTGGAGGCAAGACAACCTCTAG | 390 |
| QY | 2958 | GGATTAAAGCTAAAGAAGCTATGGACAAGGTAGTTTTTAAAGAACATA | 3005 |
| Db | 391 | GTTATCAAGCTAAAGAAGCTATGGATAAGGGAGAGCTGTTTTCAGATA | 438 |

RESULT 7

| | |
|----------|---|
| AAC38309 | |
| ID | AAC38309 standard; DNA; 469 BP. |
| XX | AAC38309; |
| XX | AC |
| XX | XX |
| DT | 17-OCT-2000 (first entry) |
| DE | Zea mays DNA fragment SEQ ID NO: 20515. |
| XX | |
| KW | Hybridisation assay; genetic mapping; gene expression control; |
| KW | protein identification; signal transduction pathway; metabolic; |
| KW | pathway; promoter; termination sequence; corn; ss. |
| XX | |
| OS | Zea mays subsp. mays. |
| XX | |
| FN | EF1033405-A2. |
| XX | |
| PD | 06-SEP-2000. |
| XX | |
| PF | 25-FEB-2000; 2000EP-0301439. |
| XX | |
| PR | 25-FEB-1999; 99US-0121825. |
| PR | 05-MAR-1999; 99US-0123180. |
| PR | 09-MAR-1999; 99US-0123548. |
| PR | 23-MAR-1999; 99US-0125788. |
| PR | 25-MAR-1999; 99US-0126264. |
| PR | 29-MAR-1999; 99US-0126785. |
| PR | 01-APR-1999; 99US-0127462. |
| PR | 06-APR-1999; 99US-0128234. |
| PR | 08-APR-1999; 99US-0128714. |
| PR | 16-APR-1999; 99US-0129845. |
| PR | 19-APR-1999; 99US-0130077. |
| PR | 21-APR-1999; 99US-0130449. |
| PR | 23-APR-1999; 99US-0130510. |
| PR | 28-APR-1999; 99US-0130891. |
| PR | 30-APR-1999; 99US-0131449. |
| PR | 30-APR-1999; 99US-0132048. |
| PR | 04-MAY-1999; 99US-0132407. |
| PR | 05-MAY-1999; 99US-0132484. |
| PR | 06-MAY-1999; 99US-0132485. |
| PR | 06-MAY-1999; 99US-0132487. |
| PR | 07-MAY-1999; 99US-0132863. |
| PR | 11-MAY-1999; 99US-0134256. |
| PR | 14-MAY-1999; 99US-0134218. |

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134376.
PR 20-MAY-1999; 99US-0134376.
PR 20-MAY-1999; 99US-0134376.
PR 21-MAY-1999; 99US-0134376.
PR 21-MAY-1999; 99US-0134376.
PR 24-MAY-1999; 99US-0134376.
PR 24-MAY-1999; 99US-0134376.
PR 25-MAY-1999; 99US-0134376.
PR 25-MAY-1999; 99US-0134376.
PR 27-MAY-1999; 99US-0134376.
PR 27-MAY-1999; 99US-0134376.
PR 28-MAY-1999; 99US-0134376.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 28-JUL-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

Db 289 TGTGAAGGCAAGCAAGCAATGATGAAGGAGAGCTT 325
RESULT 11
AAC35687
ID AAC35687 standard; DNA; 611 BP.
XX
AC AAC35687;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 11067.
XX
KW Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 102.6; DB 21; Length 1017;
Best Local Similarity 60.6%; Pred. No. 3e-13;
Matches 168; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

| | | | | |
|-----------|--|---|------|---------------|
| QY | 2719 | CATAGCATCTGGCAGAGCTTGAATATTGATCTAGAAATGGTAAGACAATCTCTCGTG | 2778 | 99US-0135124. |
| Db | 67 | CATGGCAACGAGCAGTGGCGCTCTGTCATATGGAAGATATTCAGACTGTTGATCTCAT | 126 | 99US-0135353. |
| QY | 2779 | TTCTAAAAAAGGTAAGACATACAAATTTGACATCCCTTTATTATTACTAAATTTT | 2838 | 99US-0135629. |
| Db | 127 | GTCTAGCTTCTTCGCGGTATGAAATGTGCCCTAAACCTGACAAACGCTCTGTTTTCAT | 186 | 99US-0136021. |
| QY | 2839 | AGGTCACCTGGCTGGGAAAGGAACACAGTCACCGCTGATTAAAGGATGAATTTTGCTT | 2898 | 99US-0136392. |
| Db | 187 | CGGTCACCTGGTTTCAGGAAAGGTACAGTCTCCGGTCATAAAGGATGAGTTTGT | 246 | 99US-0136782. |
| QY | 2899 | GTGCCATTTAGCCACTGGTGATATGTTGAGGCTGCAGTGGCTGCTAAACTCCACTTGG | 2958 | 99US-0137222. |
| Db | 247 | GTGTCATTTGTCTACTGGTGACATGCTTAGAGCTGCTGTGCTGCTAAGACTCCTCTTGG | 306 | 99US-0137528. |
| QY | 2959 | GATTAGGCTAAGAGCTATGCACAGTAGTATTTT | 2995 | 99US-0137502. |
| Db | 307 | TGTGAAGGCAAGGAAGCAATGGATAAGGGAGCTT | 343 | 99US-0137724. |
| RESULT 13 | | | | 99US-0138094. |
| AAC37085 | | | | 99US-0138540. |
| ID | AAC37085 standard; DNA; 1023 BP. | | | 99US-0138847. |
| AC | AAC37085; | | | 99US-0139119. |
| XX | | | | 99US-0139452. |
| DT | 17-OCT-2000 (first entry) | | | 99US-0139453. |
| DE | Arabidopsis thaliana DNA fragment SEQ ID NO: 16124. | | | 99US-0139454. |
| XX | | | | 99US-0139455. |
| KW | Hybridisation assay; genetic mapping; gene expression control; | | | 99US-0139456. |
| KW | protein identification; signal transduction pathway; | | | 99US-0139457. |
| KW | metabolic pathway; promoter; termination sequence; ss. | | | 99US-0139458. |
| OS | Arabidopsis thaliana. | | | 99US-0139459. |
| PN | EP1033405-A2. | | | 99US-0139460. |
| XX | | | | 99US-0139461. |
| PD | 06-SEP-2000. | | | 99US-0139462. |
| PF | 25-FEB-2000; 2000EP-0301439. | | | 99US-0139463. |
| PR | 25-FEB-1999; 99US-0121825. | | | 99US-0139464. |
| PR | 05-MAR-1999; 99US-0123180. | | | 99US-0139465. |
| PR | 09-MAR-1999; 99US-0123548. | | | 99US-014005. |
| PR | 23-MAR-1999; 99US-0125788. | | | 99US-014008. |
| PR | 25-MAR-1999; 99US-0126264. | | | 99US-014086. |
| PR | 29-MAR-1999; 99US-0126785. | | | 99US-014325. |
| PR | 01-APR-1999; 99US-0127462. | | | 99US-0144331. |
| PR | 06-APR-1999; 99US-0128234. | | | 99US-0144332. |
| PR | 08-APR-1999; 99US-0128714. | | | 99US-0144333. |
| PR | 16-APR-1999; 99US-0129845. | | | 99US-0144334. |
| PR | 19-APR-1999; 99US-0130077. | | | 99US-0144335. |
| PR | 21-APR-1999; 99US-0130449. | | | 99US-0144352. |
| PR | 23-APR-1999; 99US-0130510. | | | 99US-0144352. |
| PR | 28-APR-1999; 99US-0131449. | | | 99US-0144884. |
| PR | 30-APR-1999; 99US-0132048. | | | 99US-0144814. |
| PR | 30-APR-1999; 99US-0132407. | | | 99US-0145086. |
| PR | 04-MAY-1999; 99US-0132484. | | | 99US-0145088. |
| PR | 05-MAY-1999; 99US-0132485. | | | 99US-0145085. |
| PR | 06-MAY-1999; 99US-0132486. | | | 99US-0145087. |
| PR | 07-MAY-1999; 99US-0132487. | | | 99US-0145089. |
| PR | 11-MAY-1999; 99US-0132863. | | | 99US-0145192. |
| PR | 14-MAY-1999; 99US-0134256. | | | 99US-0145145. |
| PR | 14-MAY-1999; 99US-0134218. | | | 99US-0145218. |
| PR | 14-MAY-1999; 99US-0134219. | | | 99US-0145224. |
| PR | 14-MAY-1999; 99US-0134221. | | | 99US-0145276. |
| PR | 14-MAY-1999; 99US-0134370. | | | 99US-0145913. |
| PR | 18-MAY-1999; 99US-0134768. | | | 99US-0145918. |
| PR | 19-MAY-1999; 99US-0134941. | | | 99US-0145919. |
| PR | | | | 99US-0145951. |

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 102.6; DB 21; Length 1023;
Best Local Similarity 60.6%; Pred No. 3e-13;
Matches 168; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 2719 CATAGCATCTGGCAGAGCTTGAATATTGATCTCTAGAAATGGAATGGAAGCAATCTTGGGTG 2778
Db 72 CATGGCAACGAGCAGTGGCGCTTCTTGATATGGAAGATATTCAGACCGTGTGATCTCAT 131
QY 2779 TTCTAAAAAAGGTAAGACATACATTTTGACATCCCTTTTATTATTACTAAATTT 2838
Db 132 GTCTGAGCTTCTTCGCGGTATGAATGTGCTCTAAACCTGACAAACGTCTCGTTTTCAT 191
QY 2839 AGGTCCACCTGGTGGGAAAGGGAACACAGTCACCGCTGATTAAAGGATGAATTTTGCTT 2898
Db 192 CGGTCCACCTGGTTTCAGGGAAGGTACACAGTCTCCGGTCATAAAGGATGAGTTTGT 251
QY 2899 GTGCCATTTAGCCACTGGTGATATGTTGAGGCTGCGAGTGGCTGTAAACTCCACTGG 2958
Db 252 GTGTCTATTTGCTTACTGGTGACATGCTTAGAGCTGCTGTGCTGTGCTAAGACTCCTCTTGG 311
QY 2959 GATTAAGGCTAAAGAGCTATGGACAAGGTAGTTT 2995
Db 312 TGTGAAGGCAAGGAAGCAATGGATAAGGAGAGCTT 348

RESULT 14
ABL72417
ID ABL72417 standard; cdna; 295 BP.
XX ABL72417;
AC ABL72417;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:1791.
XX
KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS Zea mays.
XX
PN US2001051335-A1.
XX
PD 13-DEC-2001.
XX
PF 16-APR-1999; 99US-0294093.
XX
PR 21-APR-1998; 98US-082567P.
XX
PR (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WFI; 2002-163647/21.
XX
PT Novel purified corn tassel-derived polynucleotide useful for
PT determining altered gene expression, to recover regulatory elements and
PT to follow inheritance of desirable characteristics through hybrid
PT breeding programs
XX
PS Claim 1; SEQ ID 1791; 201pp; English.
XX
CC The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC

CC selected from those given in ABL70627 to ABL76833. The cdps sequences
 CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (I) are also useful
 CC in the evaluation, and alteration of desired characteristics associated
 CC with growth and development, disease resistance, environmental
 CC adaptability, quality and yield, and as molecular markers for studying
 CC inheritance of multigene traits in a plant breeding program. (I) can be
 CC used to produce a tassel-specific profile of gene transcription, a
 CC transcript image, to clone regulatory elements for use in transformation
 CC vectors, to express a polypeptide, to identify, isolate or extend
 CC identical or related corn tassel nucleic acid sequences from DNA
 CC libraries, in nucleic acid hybridisation or amplification technologies,
 CC as query sequences to determine homology of known sequences, as probe
 CC for use in Southern or Northern hybridisation, and to identify the
 CC presence of and/or to determine the degree of similarity between two
 CC (or more) nucleic acid sequences.
 XX
 SQ Sequence 295 BP; 69 A; 78 C; 80 G; 68 T; 0 other;

Query Match 2.4%; Score 102.2; DB 24; Length 295;
 Best Local Similarity 84.1%; Pred. No. 2.2e-13;
 Matches 127; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
 QY 2838 TAGTCCACCTGGCTCGGAAAGGGAACACAGTCACCGCTGATTAAGGATGAATT-TTGC 2896
 DB 144 TCGGCCACCTGGCTCGGAAAGGGAACAGTCACCTCCCTTATTAAAGGATGAATTGC 203
 QY 2897 TTGTGCATTTAGCCACTGGTATATGTTGAGGGCTGCAGTGGCTGCTAAACTCCACTT 2956
 DB 204 CTGTGCCATTTAGCCACTGGTATATGTTGAGGGCTGTGGCAGCCCAAGCACTCTA 263
 QY 2957 GGATTAAGGCTAAAGAACTATGGCAAGG 2987
 DB 264 GGTATCAAGCTAAAGAACTATGGTAAGG 294

RESULT 15
 ABN98716/c
 ID ABN98716 standard; DNA; 894 BP.
 XX
 AC ABN98716;
 XX
 XX 01-AUG-2002 (first entry)
 XX
 DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 484.
 XX
 KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
 KW disease; crop; thale cress; tolerance factor; insect; pathogen;
 KW nutrition; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 XX US2002023281-A1.
 XX
 XX 21-FEB-2002.
 XX
 XX 26-JAN-2001; 2001US-0770445.
 XX
 XX 27-JAN-2000; 2000US-178472P.
 XX
 XX (GORLACH J.
 PA (ANY)/ AN Y.
 PA (HAMI)/ HAMILTON C M.
 PA (PRIC)/ PRICE J L.
 PA (RAIN)/ RAINES T M.
 PA (YUYU)/ YU Y.
 PA (RAME)/ RAMEKA J G.
 PA (PAGE)/ PAGE A.
 PA (MATH)/ MATHAW A V.
 PA (LEDF)/ LEDFORD B L.
 PA (WOES)/ WOESSNER J P.

PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRIC/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX
 DR WPI; 2002-400781/43.
 XX
 PT New Arabidopsis thaliana nucleic acid for identifying homologous genes,
 PT producing compositions that modulate the expression or function of its
 PT encoded protein, and mapping functional regions of protein -
 XX
 PS Claim 1; SEQ ID NO 484; 49pp + Sequence Listing; English.
 XX
 CC The invention relates to an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridising under stringent conditions
 CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
 CC given in the specification or its fragment. A polypeptide (II) encoded by
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
 CC useful for screening a candidate agent for its biological effect. (I) is
 CC useful in identifying homologous or related genes, in producing
 CC compositions that modulate the expression or function of its encoded
 CC protein, mapping functional regions of the protein and in studying
 CC associated physiological pathways. (I) is also useful for the genetic
 CC manipulation of cells, particularly plant cells. (I) is also useful in
 CC screening assays of various plant strains to determine the strains that
 CC are best capable of withstanding a particular disease or environmental
 CC stress. (II) and (III) are useful for screening of biologically active
 CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
 CC pathways. The screened agents are useful in improved methods of treating
 CC crops to prevent or treat disease. (II) are also useful in screening
 CC programs to identify agents that mimic or enhance the action of tolerance
 CC factors. Such agents are useful in improved methods of treating crops to
 CC enhance their tolerance to environmental stress. (I) is also useful
 CC for enhancing or inhibiting production of a biosynthetic product in a
 CC plant. (III) is useful for identifying other mediators that may induce
 CC expression of proteins of interest, for establishing the extent to which
 CC any specific insect and/or pathogen is responsible for damage to a
 CC particular plant, for identifying other mediators that enhance or induce
 CC tolerance to environmental stress, for identifying factors involved in
 CC biosynthetic pathways of nutritional, commercial, or medicinal value and
 CC for identifying productions of nutritional, commercial or medicinal
 CC value. (IV) is useful in the study of genetic function and regulation,
 CC for alteration of the cellular metabolism and for screening compounds
 CC that may affect the biological function of the gene or gene products.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=999909770445.
 XX
 SQ Sequence 894 BP; 246 A; 221 C; 155 G; 258 T; 14 other;

Query Match 2.3%; Score 101.8; DB 24; Length 894;
 Best Local Similarity 71.9%; Pred. No. 4.3e-13;
 Matches 133; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 3303 AAATGGTCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCTAA 3362
 DB 583 AAGCAGAGAAGCTTGATGAGATGCTTAATAGAGGGAGCTCAGATAGATAAGGTGCTTA 524
 QY 3363 ATTTTCGAATTGATGCAATCTGGAGACGATACCGTCTGTTGGATCCACCAT 3422
 DB 523 ATTTTCGATCGATGATTCGGTTCTCGAAGAAAGAAATTAAGGAGGTGATTCACCTT 464
 QY 3423 CAAGTGGTAGATCTTATCATACAAAATTTGCTCTCTCTTAAGACTCTCGGACTTGTATGATG 3482

|||||
Db 463 CAAAGTGGAGAGAGCTATCATCTAAATTCGCACCTCCTAAAGTTCAGGAGTCGATGATC 404
Qy 3483 TAACT 3487
Db 403 TGAAT 399

Search completed: March 15, 2003, 10:55:20
Job time : 843.479 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:55 ; Search time 137.416 Seconds
(without alignments)
9687.969 Million cell updates/sec

Title: US-09-802-937-1

Perfect score: 4341

Sequence: 1 ctgcaggaagattaattagg.....gtatatcttcattgttttt 4341

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/FCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| C 1 | 57.2 | 1.3 | 7218 | 1 US-08-232-463-14 | Sequence 14, Appl |
| C 2 | 57 | 1.3 | 11831 | 4 US-08-961-527-65 | Sequence 65, Appl |
| C 3 | 53 | 1.2 | 7218 | 1 US-08-232-463-14 | Sequence 14, Appl |
| C 4 | 50.2 | 1.2 | 6124 | 4 US-08-213-419B-3 | Sequence 3, Appl |
| C 5 | 50 | 1.2 | 19124 | 2 US-08-487-826B-13 | Sequence 13, Appl |
| C 6 | 49.2 | 1.1 | 669 | 4 US-09-134-001C-452 | Sequence 452, Appl |
| C 7 | 49.2 | 1.1 | 19124 | 2 US-08-487-826B-13 | Sequence 13, Appl |
| C 8 | 46.8 | 1.1 | 636 | 4 US-08-998-416-1137 | Sequence 1137, Ap |
| C 9 | 46.6 | 1.1 | 1850 | 3 US-08-617-860B-32 | Sequence 32, Appl |
| C 10 | 46.6 | 1.1 | 2389 | 1 US-08-153-563-2 | Sequence 2, Appl |
| C 11 | 46.6 | 1.1 | 2389 | 2 US-08-460-507-2 | Sequence 2, Appl |
| C 12 | 46.6 | 1.1 | 4098 | 2 US-08-605-106-4 | Sequence 4, Appl |
| C 13 | 46.6 | 1.1 | 4276 | 1 US-07-973-324A-3 | Sequence 3, Appl |
| C 14 | 46.6 | 1.1 | 4276 | 1 US-08-343-380-3 | Sequence 3, Appl |
| C 15 | 46.6 | 1.1 | 4276 | 4 US-09-072-435-3 | Sequence 3, Appl |
| C 16 | 46.6 | 1.1 | 4276 | 4 US-09-072-917A-3 | Sequence 3, Appl |
| C 17 | 46.2 | 1.1 | 2304 | 1 US-08-464-266-1 | Sequence 1, Appl |
| C 18 | 46.2 | 1.1 | 2304 | 1 US-08-464-272-1 | Sequence 1, Appl |
| C 19 | 46.2 | 1.1 | 2304 | 4 US-08-464-514-1 | Sequence 1, Appl |
| C 20 | 46.2 | 1.1 | 2304 | 4 US-08-486-403-1 | Sequence 1, Appl |
| C 21 | 45.8 | 1.1 | 5852 | 1 US-07-867-106-2 | Sequence 2, Appl |
| C 22 | 45.8 | 1.1 | 9048 | 3 US-08-973-273-4 | Sequence 4, Appl |
| C 23 | 45.4 | 1.0 | 854 | 2 US-08-829-027-2 | Sequence 2, Appl |
| C 24 | 45.4 | 1.0 | 854 | 2 US-09-225-366-2 | Sequence 2, Appl |
| C 25 | 45.4 | 1.0 | 1751 | 4 US-09-149-476-68 | Sequence 68, Appl |
| C 26 | 45 | 1.0 | 636 | 4 US-08-998-416-1137 | Sequence 1137, Ap |
| C 27 | 44.2 | 1.0 | 837 | 4 US-08-998-416-288 | Sequence 288, Appl |

Sequence 595, Appl
Sequence 33, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 651, Appl
Sequence 3, Appl
Sequence 36, Appl
Sequence 786, Appl
Sequence 76, Appl
Sequence 76, Appl
Sequence 76, Appl
Sequence 2, Appl
Sequence 14, Appl
Sequence 1, Appl
Sequence 36, Appl
Sequence 35, Appl
Sequence 2202, Ap

ALIGNMENTS

RESULT 1

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 1.3%; Score 57.2; DB 1; Length 7218;


```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match
Best Local Similarity 1.2%; Score 50; DB 2; Length 19124;
Matches 273; Conservative 0; Mismatches 325; Indels 8; Gaps 2;

QY 483 CAATTCAACCGTTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAAAAT 542
DB 15593 CAAAAAATAATTAATAAATAATTTTATATAAATAAATAAATGATTATAAATAAATAAATAA 15652
QY 543 TAATATCATATAAATAAAGAGTTGTTAATTAACAACCAATGTTTCTGCTACTACAACCTCTAA 602
DB 15653 CAAAAGAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15712
QY 603 TTTGTAAATCTTATTTTCAGTCACAAAATCCAAATTTCCAAATTAAGAAAAATAAAGCTAG 662
DB 15713 AAGAATAAATAATATATAAATAAATAATATATATCAATAAATAAATAAATAAATAA 15771
QY 663 ACGGCTAAGCCCAACCCATCTAAGGCTTAAGTTTCGAGAGGTGAAGTACGACGAAAAATATG 722
DB 15772 ATGTTAAAAAATAATATATACATAAATAAATAAATAAATAAATAAATAAATAAATAA 15831
QY 723 ATGGTTTATATATAGATTTTTTTTAAATAAATCTTCACATAAATTTCTTTAGGAACATA 782
DB 15832 TAAATAAATAAATAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15891
QY 783 TCATTTAATGTTTGAATAAATCGTCACATAAAGAACTAAGAACGATGATGGGAACA 842
DB 15892 AAAAAAATAAATAAATTTAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15951
QY 843 AGAGAAAAACAGCCCTTAAGGCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAACGC 902
DB 15952 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16004
QY 903 ATGATAAAGAGAAAGCTCATAGACATATTTACTAGATATTTATATATATATATATATAT 962
DB 16005 ACCATAACTACATAACAACATTTTACACATACATATATATATATATATATATATATAT 16064
QY 963 AAAAAAATAATTTTGAATTTTTTAAACAATGATGCATAAATATTTTTTAAAAAACAC 1022
DB 16065 CATACATTTTACACATACATATATATATATATATATATATATATATATATATATATAT 16124
QY 1023 ACCAATTTAACCTTTAAAGAGATCCTTAATAGGAAACGAGGAAGTTAAAGATTACCGA 1082
DB 16125 ATACATATATACATTAAACAACACATATATAATACCTAAATACATATATATATATATAT 16184
QY 1083 AGTGT 1088
DB 16185 TATGTT 16190

RESULT 6
US-09-134-001C-452
; Sequence 452, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 452
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-452

Query Match
Best Local Similarity 1.1%; Score 49.2; DB 4; Length 669;
Matches 99; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 3318 ATGAAATGTGGCCAAACAAAGTACTAAGATTGACAAGGTTCTAAATTTTGAATTTGATG 3377
DB 3311 ATCAAATTTATGCTGAGCTTGATAGAGAAATGATGCTGTCTAATTAATATCGAAGTTCTG 370
QY 3378 ATGCAATATCTGAAGAAAGAAATACCGGTCGTTGGATCCACCCATCAAGTGGTATCTT 3437
DB 371 AGGAAGAATTAATGAATCGTCTTACAGGTCGTCGATCTGTGAGAAATGTGGTACAACAT 430
QY 3438 ATCATACAAAAATTTGCTCTCTTAAGACTCTCGGACTTGTAGTGTAAAGTCTATACCAGAT 3497
DB 431 ATCATCTTGTATTTAATCTCTCAAAGGTTGATGTTATATGATATCGATCGTGGAAAGT 490
QY 3498 TA 3499
DB 491 TA 492

RESULT 7
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

US-08-487-826B-13

```
Query Match          1.1%; Score 49.2; DB 2; Length 19124;
Best Local Similarity 47.3%; Pred. No. 0.032;
Matches 244; Conservative 0; Mismatches 268; Indels 4; Gaps 3;

QY 3604 TTATACATATTTCCCTGCTCATCATCGATCATCGTAGTACTCTGCTTACTTGTGTTTTT 3663
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15945 TTTAAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15886

QY 3664 AAGAACAAACATGAGCCATCTCTTCCAAATATAAAGAGTTTCATTTGAGTACCTCC 3723
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15885 AATTAATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 15826

QY 3724 TCTTGCCACTGTTTAAATAGTTTGGCTTCCAAATAGTTAAATAGGTAGTGTGATATGGA 3783
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15825 TTTTATTTTAAATAAATTTTATTTTATTTATTTATGTATATTTTATTTTAAATTTT 15766

QY 3784 TATATTTTCTGTTGGTGTGTTCTCGTACGAAGAGTAAA--ATGCACCTTTTGTGGA 3841
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15765 AATTTTATTTTATTTATGATATATATTTTATTTTAAATATATTTTCTTTTCTTTT 15706

QY 3842 CAGAATATGATAGCGATGTTTGCATGACACACTTTTGTCTTCTTCTGACATATATGC 3901
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15705 GTTTTATGATATATATTTTATTTTATTTTATTTTATTTTCTTTTCTTTTAT 15646

QY 3902 CTGTTTAAAGTGGCCATAAATAGATACATGACATGTTTTTTTGTAGCAGGAGATTGTATAT 3961
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15645 TTTTATTAATCATTTTATTTTATAT-AAAAATTTTATTTTATTTTATTTTATTAATC 15587

QY 3962 TGTTTCTATTTCTCCAAATPAAAGCATATCTTCTTTAGCAATGATTCATGTGGACA 4021
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15586 TTTTTCATTTTATTTATCTATCAAAATTTATATTTTATTTATTAATTTTATTTTAA 15527

QY 4022 TATTTGTGCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTAAATGGTAATA 4081
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15526 AAAATTTTCTCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 15468

QY 4082 TTATGTCACCTTCGCTCCCTGATTTGCTTGTCTTT 4117
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15467 CATTTTCTTTTTCATTTTAAATAAATTTGTTTTT 15432
```

```
RESULT 8
US-08-998-416-1137/c
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; NUMBER OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1692RP
; US-08-998-416-1137
```

```
Query Match          1.1%; Score 46.8; DB 4; Length 636;
Best Local Similarity 44.6%; Pred. No. 0.021;
Matches 274; Conservative 0; Mismatches 332; Indels 8; Gaps 2;

QY 446 TTCAAAAAGAAATTAATTTTTCATGAAACGCAATTCACCGTTGAGAAATAT 505
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 636 TTTTATAGTATTTTAACTACATTCITTATATATTTTATTAATTAATTAATGAT 577

QY 506 GCTGTCTATAAATAGTAGTCTAGTGCAGAAACAAAATTAATATCACATAAAAAAGAGGT 565
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 576 AAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 517

QY 566 TCTTAATTTACAAACCATGTTTCGTACTACACTCTAATTTGTAATTTCTTATTTTCAGTCA 625
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 516 TATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 457

QY 626 CAAAATTCCAATTTCCAAATTAAGAAAAATAAACGTAGACGCGCTAAGCCCAACCTCTAAG 685
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 456 TCTTTATAAAAAAGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 397

QY 686 GCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATATGATGGTTTATTA-----ATATG 738
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 396 ATAAAAATAAATAATTTACATAATTTTAAATAATTAATTAATTAATTAATTAATTAAT 337

QY 739 ATTTTATTTTAAATAACTTTTCACATAAAATTTCTTTAGGAAACATATCATTTTAATGGTTGA 798
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 336 ATTTTAAATAACAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 277

QY 799 AAAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCC 858
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 276 TAAAGAAAAATAATAATCTAATAATAATTTTAAATACTAATTTTAAATACTGACATAGAC 217

QY 859 TTAAGGCTTTTGATCCCTCTAGTTGGAGGTGATTTTCAAACGCGATGATAAACGAGAAAG 918
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 216 TAAATAGTATTCATATTAATAATAATAATAATAATAATAATAATAATAATAATAATGAT 157

QY 919 CTCATTAGCACATTTACTTAGATATTTTATAATTTATAAACTTGAAAAAATAATTTATTT 978
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 TAAATAGTATTTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 98

QY 979 GAATTTTAAACAAATGTATGCATAAAATTTTATTTTAAACAAACACACCAATTTAACCCTTT 1038
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 AATAATGATAATAATAGTTTAAATAATAATAATAATAATAATAATAATAATAATAATAAT 38

QY 1039 AAAAAGCATCTCTAA 1052
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 ATTAATCTTTATAA 24
```

RESULT 9
US-08-617-860B-32/c

Db 2075 AACATGAATACCTGATCGCACACAGAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2016
Qy 899 ACCCATGATAAAGAGAAAGCTCATTAGCACATTTACTTAGATATTTATTAATTAATAA 958
Db 2015 ATGTAGCGAAACGAGCTGAGCCATTAGATATAATTAATGAGTATTAATCTATTTTAA 1956
Qy 959 CTTGAAAAAATATTTATTTGAAATTTTAAACAATGTATGCATATAATTTTAAAAA 1018
Db 1955 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCAAAA 1896
Qy 1019 ACACACCAATTTAACCCCTTTAAAA 1043
Db 1895 ATCATACCGTTTAGTAGTTCAGGAA 1871

RESULT 11
US-08-460-507-2/c
; Sequence 2, Application US/08460507
; Patent No. 5984628
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: PROCESS FOR PROTEIN PRODUCTION
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460.507
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,563
; FILING DATE: 16-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION/DOCKET NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 2000-0452.41
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2389
; OTHER INFORMATION: /standard_name="RAMV-1A"

Query Match 1.1%; Score 46.6; DB 2; Length 2389;
Best Local Similarity 51.7%; Pred. No. 0.049;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 839 AACAAAGAAAAACAGAGCTTAAAGCTTCTTGATCCTCTAGTTGGAGGTTGATTTCAA 898
Db 2075 AACATGAATACCTGATCGCACACAGAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2016
Qy 899 ACCCATGATAAAGAGAAAGCTCATTAGCACATTTACTTAGATATTTATTAATTAATAA 958

Db 2015 ATGTACCAAAACGAGCTGAGCCATTAACTAATTAATGAGTATTAATCTATTTTAA 1956
Qy 959 CTTGAAAAAATATTTATTTGAAATTTTAAACAATGTATGCATATAATTTTAAAAA 1018
Db 1955 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCAAAA 1896
Qy 1019 ACACACCAATTTAACCCCTTTAAAA 1043
Db 1895 ATCATACCGTTTAGTAGTTCAGGAA 1871

RESULT 12
US-08-605-106-4/c
; Sequence 4, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235.001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4098 Base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: : DNS (genomic)
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: CITEg1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
; LOCATION: ..3303, 3391..3459, 3672..3941)
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 1797..1799
; FEATURE:
; NAME/KEY: exon II
; LOCATION: 1787..2294

QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
DB 1889 ATCATACCGTTTGTAGTTCAGGAA 1865

RESULT 14
US-08-343-380-3/c
; Sequence 3, Application US/08343380
; Patent No. 5712112
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Liu, Li-Fei
; TITLE OF INVENTION: Gene Expression System Comprising the
; TITLE OF INVENTION: Promoter Region of the Alpha-Amylase Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343.380
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/973,324
; FILING DATE: 04-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 31149
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Rice (Oryzae sativa)
; STRAIN: CV. M202
; IMMEDIATE SOURCE:
; LIBRARY: (EMBL) genomic
; CLONE: '-Amy7-C
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: 3952)
; PUBLICATION INFORMATION:
; AUTHORS: Yu et al., Su-May
; TITLE: Regulation of '-amylase-encoding gene expression
; TITLE: in germinating seeds and cultured cells of rice
; JOURNAL: Gene
; VOLUME: in press
; US-08-343-380-3

Query Match 1.1%; Score 46.6; DB 1; Length 4276;

Best Local Similarity 51.7%; Pred. No. 0.066;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAGAAAAACACAGCCTTAAGCCTTCTTGATCCTCTAGTTGGAGTTGATTTCAA 898
DB 2069 AACATGAATACCTGATCGCACACAGAGCATTTGAGACAGAGGAAAGATCGATTGAGA 2010
QY 899 ACGCATGATAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTTATAATATAAA 958
DB 2009 ATGTACGCAAAACGAGCTGAGCCATTAACTGATATAATTGAGTATTAACTATTTAAA 1950
QY 959 CTTGAAAAAATATTATTGTTGATTTTAAACATGATGCATAAATATTATTTAAAA 1018
DB 1949 TTTTAAAAATAGATTATAATATTTTAAAGTAACTTCTCTATAGAAAATTTTGCAAAA 1890
QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
DB 1889 ATCATACCGTTTGTAGTTCAGGAA 1865

RESULT 15
US-09-072-435-3/c
; Sequence 3, Application US/09072435
; Patent No. 6215051
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Liu, Li-Fei
; APPLICANT: Chan, Ming-Tsair
; TITLE OF INVENTION: GENE EXPRESSION SYSTEM COMPRISING THE
; TITLE OF INVENTION: PROMOTER REGION OF THE ALPHA-AMYLASE GENES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,435
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,792
; FILING DATE: 29-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/973,324
; FILING DATE: 04-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28123/34274
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Rice (Oryzae sativa)
; STRAIN: CV. M202
; IMMEDIATE SOURCE:
; LIBRARY: (EMBL) genomic
; CLONE: '-Amy7-C

```
;
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
;
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
;
; PUBLICATION INFORMATION:
; AUTHORS: Yu et al., Su-May
; TITLE: Regulation of '-amylase-encoding gene expression
; TITLE: in germinating seeds and cultured cells of rice
; JOURNAL: Gene
; VOLUME: in press
;
; US-09-072-433-3
;
Query Match      1.1%; Score 46.6; DB 4; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.066;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAAAGAGAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTTGAGGTTGATTTCAA 898
Db 2069 ACATGAATACCTGATCGCACAGAGCATTTGAGAGCAGAGAAAAGAAATCGATTGAGA 2010

QY 899 ACGCATGATAACGAGAAAAGCTCATTAGCACATTATTACTTAGATATTTTATAATTATAAA 958
Db 2009 ATGTACGCAAAACGAGCTGAGCCATTAAACGTATAATTAAATTGAGTATTAACTATTTTAAA 1950

QY 959 CTTGAAAAAATATTTATTGAAATTTTAAACAATGTATGCATAAATTTTAAAAA 1018
Db 1949 TTTTAAAAATAGATTAATATAATTTTAAAGTAACTTCTCTATAGAAAATTTTGCAAAAA 1890

QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
Db 1889 ATCATACCGTTTAGTAGTTCAGGAA 1865

Search completed: March 15, 2003, 16:59:03
Job time : 365.416 secs
```

| Result No. | Query | | | DB | ID | Description |
|------------|--------|-------|--------|----|---------------------|-------------------|
| | score | Match | Length | | | |
| 1 | 3987.8 | 31.9 | 4354 | 10 | US-09-802-927-1 | Sequence 1, Appli |
| 2 | 1384.6 | 91.9 | 1443 | 10 | US-09-802-927-2 | Sequence 2, Appli |
| 3 | 122 | 2.8 | 404 | 10 | US-09-878-574-3892 | Sequence 3892, Ap |
| C 4 | 112.8 | 2.6 | 985 | 10 | US-09-770-445-271 | Sequence 271, App |
| 5 | 102.2 | 2.4 | 295 | 10 | US-09-294-093B-1791 | Sequence 1791, Ap |
| C 6 | 101.8 | 2.3 | 894 | 10 | US-09-770-445-484 | Sequence 484, App |
| 7 | 92.8 | 2.1 | 292 | 10 | US-09-923-876-3999 | Sequence 3999, Ap |
| 8 | 92 | 2.1 | 267 | 10 | US-09-923-876-6281 | Sequence 6281, Ap |
| C 9 | 62.4 | 1.4 | 2000 | 10 | US-09-887-576-828 | Sequence 828, App |
| 10 | 61.6 | 1.4 | 645 | 10 | US-09-815-242-6911 | Sequence 6911, Ap |
| 11 | 57 | 1.3 | 639 | 10 | US-09-815-242-9121 | Sequence 9121, Ap |
| 12 | 57 | 1.3 | 908 | 12 | US-10-044-090-213 | Sequence 213, App |
| C 13 | 55.8 | 1.3 | 8416 | 8 | US-08-910-386A-4 | Sequence 4, Appli |
| C 14 | 55.8 | 1.3 | 13341 | 8 | US-08-910-386A-1 | Sequence 1, Appli |
| C 15 | 55.8 | 1.3 | 19639 | 8 | US-08-910-386A-6 | Sequence 6, Appli |
| C 16 | 55.6 | 1.3 | 355 | 10 | US-09-860-352-11750 | Sequence 11750, A |
| 17 | 53.6 | 1.2 | 398 | 10 | US-09-960-352-11004 | Sequence 11004, A |
| C 18 | 53.2 | 1.2 | 5940 | 8 | US-08-910-386A-10 | Sequence 10, Appl |
| C 19 | 52.2 | 1.2 | 238 | 10 | US-09-864-761-31345 | Sequence 31345, A |

Qy 181 TCTTGTAACCTGTTTGAATACATAGACAGTGCCAAAC-GTTTGTATGGCTCCTATTG 239
Db 181 TCTTGTAACCTGTTTGAATACATAGACAGTGCCAAACGTTTGTATGGCTCCTATTG 240
Qy 240 GCTCGTGATGACTGCTGTGCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 299
Db 241 GCTCGTGATGACTGCTGTGCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 300
Qy 300 ACCGAAACCCCAAGATATTCTATTCCACCTCAGGGTAATGTGCTGAACTATGCAAT 359
Db 301 ACCGAAACCCCAAGATATTCTATTCCACCTCAGGGTAATGTGCTGAACTATGCAAT 360
Qy 360 GAATACAAATCCCAAAATATCATGTTATCTATCTTCTCAAAATGAAATTTGAGTCCA 419
Db 361 GAATACAAATCCCAAAATATCATGTTATCTATCTTCTCAAAATGAAATTTGAGTCCA 420
Qy 420 ACTGAGACTGCAATACGATTTTTCTTTTCAAAAAGAAATATTAAATTTTTTTTTCATGAA 479
Db 421 ACTGAGACTGCAATACGATTTTTCTTTTCAAAAAGAAATATTAAATTTTTTTTTCATGAA 480
Qy 480 ACCGAATTCACCGTTCGAAATATGCTGTCATAATAGTAGTCTAGTCAGAGAAACA 539
Db 481 ACCGAATTCACCGTTCGAAATATGCTGTCATAATAGTAGTCTAGTCAGAGAAACA 540
Qy 540 AATTAAATACACATATAAAAGAGGTGTTAAATTACAAACCATGTTTCGTACTACAACCTC 599
Db 541 AATTAAATACACATATAAAAGAGGTGTTAAATTACAAACCATGTTTCGTACTACAACCTC 600
Qy 600 TAATTTGTAATTTAATTTTCACTCAAAATTTCCAAATTTCCAAATTTAAGAAATAAAG 659
Db 601 TAATTTGTAATTTAATTTTCACTCAAAATTTCCAAATTTCCAAATTTAAGAAATAAAG 660
Qy 660 TAGACGCTAAGCCCAACCCATCTAAGCTTAAGTTTCGAGAGGTGAAGTACGACGAAAT 719
Db 661 TAGACGCTAAGCCCAACCCATCTAAGCTTAAGTTTCGAGAGGTGAAGTACGACGAAAT 720
Qy 720 ATGATGTTTATTAATATGATTTTTTTTTTAATAACTTTTCACTCAAAATTTTCTTTAGGAAAC 779
Db 721 ATGATGTTTATTAATATGATTTTTTTTTTAATAACTTTTCACTCAAAATTTTCTTTAGGAAAC 780
Qy 780 ATATCATTTAATGTTTGAACACGTGCACATAAGAAACTAAGACGATGAGTTGGGAA 839
Db 781 ATATCATTTAATGTTTGAACACGTGCACATA--GAAACTAAGACGATGAGTTGGGAA 838
Qy 840 ACAAGAGAAAAACACACCTTAAGGCTTCTGTATCTCTAGTTGGAGGTGATTTTCAAA 899
Db 839 ACAAGAGAAAAACACACCTTAAGGCTTCTGTATCTCTAGTTGGAGGTGATTTTCAAA 898
Qy 900 CGCATGATAACGAGAAAGCTCATAGCACATTTACTTAGATATTTAATATTAAC 959
Db 899 CGCATGATAACGAGAAAGCTCATAGCACATTTACTTAGATATTTAATATTAAC 958
Qy 960 TTGAAAAAATATTATTGTAATTTTTTAAACAATGATGCATAATTTATTTTTTAAAAA 1019
Db 959 TTGAAAAAATATTATTGTAATTTTTTAAACAATGATGCATAATTTATTTTTTAAAAA 1018
Qy 1020 CACACCAATTTAACCTTTTAAAGCATCTTAATAGGAAACGAGGAGTTAAGATTTAC 1079
Db 1019 CACACCAATTTAACCTTTTAAAGCATCTTAATAGGAAACGAGGAGTTAAGATTTAC 1078
Qy 1080 CGAA--GTGTTTGGATAATGAAAAATGGGGTGGGATTAGAATTTGTAATGAATCAGG 1136
Db 1079 CGAAGGTGCTTGGATAATGAAAAATGGGGTGGGATTAGAATTTGTAATGAATCAGG 1138
Qy 1137 TTAGGATTAATAATTTAAATGAAAGGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTT 1196
Db 1139 TTAGGATTAATAATTTAAATGAAAGGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTT 1198
Qy 1197 TTTGGGGTAGAAAAATTATTTGCCATACATCCCGAGAGCGGTGCTGCTGGTGGG 1256
Db 1199 TTTGGGGTAGAAAAATTATTTGCCATACATCCCGAGAGCGGTGCTGCTGGTGGG 1258
Qy 1257 CAGAAAGCTCTTTTTCGTTGGAAAAAACAACCTGCTTTAAAAAGGAAACAGAAAGGCCAG 1316

Db 1259 CAGAAAGCTCTTTTTCGTTGGAAAAAACAACCTGCTTTAAAAAGGAAACAGAAAGGCCAG 1318
Qy 1317 CTTT-GTTTGTCACCGTCTCACAGAAACGAAACAAAGAGCCCAACACCTAAACCTCTCTC 1375
Db 1319 CTTTGGTTGTACCGTCTCACAGAAACGAAACAAAGAGCCCAACACCTAAACCTCTCTC 1378
Qy 1376 GATCCGACCGAGACTCCTCCATTTTCAGCGCGCAACGCGAGAGCAACGCAACGCGAGTTGG 1435
Db 1379 GATCCGACCGAGACTCCTCCATTTTCAGCGCGCAACGCGAGAGCAACGCGAGTTGG 1438
Qy 1436 TCGACGAAACAAAGCTAGTCAGTAGTTGTTGCTGCGAGATGCGCGCAACCTCGAGGAC 1495
Db 1439 TCGACGAAACAAAGCTAGTCAGTAGTTGTTGCTGCGAGATGCGCGCAACCTCGAGGAC 1498
Qy 1496 GTCCGCTCGATGAGCTGATGACGAGCTGCTCCGCGCATGAAGTGCAGCTCCAAGGCC 1555
Db 1499 GTCCGCTCGATGAGCTGATGACGAGCTGCTCCGCGCATGAAGTGCAGCTCCAAGGCC 1558
Qy 1556 GACAAAGCGCTCATCTCGTTCGCTAACCGCGCTCTTCTTCTTCTCTCTCTCTCTCTCT 1615
Db 1559 GACAAAGCGCTCATCTCGTTCGCTAACCGCGCTCTTCTTCTTCTCTCTCTCTCTCTCT 1618
Qy 1616 CTCTCTCGTGGTTTGGTTTGGTTGGTTCGCTCCGCTTCGCGCGCGCTGCTGCCCTG 1675
Db 1619 CTCTCTCGTGGTTTGGTTTGGTTGGTTCGCTCCGCTTCGCGCGCGCTGCTGCCCTG 1678
Qy 1676 GTCCGAGATGCGGAAAGATTCGCTGTGGCTTCGCTGGGATTTGCGTTAGATCCGTGCA 1735
Db 1679 GTCCGAGATGCGGAAAGATTCGCTGTGGCTTCGCTGGGATTTGCGTTAGATCCGTGCA 1738
Qy 1736 AGCTGTCTCCCTTCTTTTGGGTCAAACTCGGTTTCTTCTTCCGCTCTCTCTCTCTG 1795
Db 1739 AGCTGTCTCCCTTCTTTTGGGTCAAACTCGGTTTCTTCTTCCGCTCTCTCTCTG 1798
Qy 1796 TTTTTCGAAACAAAAATGTTTGGCTTCGAGGAAATCTTGTAATAATTTACTTGTGTTT 1855
Db 1799 TTTTTCGAAACAAAAATGTTTGGCTTCGAGGAAATCTTGTAATAATTTACTTGTGTTT 1858
Qy 1856 CCACCTTTGAAACTCGGCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1915
Db 1859 CCACCTTTGAAACTCGGCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1917
Qy 1916 CGATGAGTCGATGCGGTTTTCATTTGACGAAAGGATTTAGTGACAG-AATCAAGC 1974
Db 1918 CGATGAGTCGATGCGGTTTTCATTTGACGAAAGGATTTAGTGACAGAAATCAAGC 1977
Qy 1975 GAAAGAAATATTATATGATCGGTTCGACGATATAGGAAATTCAGGCGTTTAAAG 2034
Db 1978 GAAAGAAATATTATATGATCGGTTCGACGATATAGGAAATTCAGGCGTTTAAAG 2037
Qy 2035 TATAGAAATTTATTCGATTTCAGTAGTATGTTGAAATTTGGAGGTTCATTTTGTGTTCT 2094
Db 2038 TATAGAAATTTATTCGATTTCAGTAGTATGTTGAAATTTGGAGGTTCATTTTGTGTTCT 2097
Qy 2095 CACACATGTTTATTTAAATTTTCATGTCGCAATGGTGTGAGGATGACGAGAAACAA 2154
Db 2098 CACACATGTTTATTTAAATTTTGTGTCAGTGCCA--TGTGTTGAGGATGAGC-AGAAAAA 2154
Qy 2155 CAGGCGGTTTCGTTGTTTTCGTCGATGCGAACTTTTTTAAATAACCAAAACAGAAA 2214
Db 2155 CAGGCGGTTTCGTTGTTTTCGTCGATGCGAACTTTTTTAAATAACCAAAACAGAAA 2214
Qy 2215 GGTGCAACTTTATGATGATAAGGTTTGGAAATGTCACCTCGGAATTAACCAAGAA 2274
Db 2215 GGTGCAACTTTATGATGATAAGGTTTGGAAATGTCACCTCGGAATTAACCAAGAA 2274
Qy 2275 GATAGACGTTTACTCTTGTTCATATTTCACTACTATAACCAAAATCGATCTGTATGTT 2334
Db 2275 GATAGACGTTTACTCTTGTTCATATTTCACTAGTAAACCAAAATCGATCTGTATGTT 2334
Qy 2335 TATGTAACCTGTGTGACAAACAAATTTCAAAATTAACCTTGTCTGTTGGCAATGCTCTG 2394

| | | | |
|----|------|--|-------|
| Db | 2335 | TATGTAAACTGTGTGACACAAACAAATTTCAAATTAACACTTGTCTGTCTGGCGCAATGCTCTG | 23394 |
| Qy | 2395 | CTTCAGCAATGCACACTACTCTAAATTTATCTGGATCATCAAAACAAATCATTTATATGAAGGTTT | 2454 |
| Db | 2395 | CTTCAGCAATGCACACTACTCTAA--TTTACTGGTCACTCAACAAACAAATCATTTATGAAGGTTT | 2452 |
| Qy | 2455 | ATTCTATTCTGTTCTTTTACGTTTCATGAAGTATAGCTAAATTTACAACTCTCTGCACCTACC | 2514 |
| Db | 2453 | ATTCTATT--GTTCTTTTACCTTCATGAAGTATAGCTAAATTTACAAACACTCTGCACCTACC | 2510 |
| Qy | 2515 | TAATTTTAGTGGACCGAAATCTAGTCATCTCTGCGCTTAGCTATCTCCAAATTCATGGCG | 2574 |
| Db | 2511 | TAATTTAAGTGGACCGAATCTAGTCATCTCTGCTCTTAGCTATCTCCAAATTCATGGCT | 2570 |
| Qy | 2575 | TGTTCTTAATTCATGCAATGTCCTAGAG--ACTGGTGTGTTTAACAGCATTTGAAGAATG | 2633 |
| Db | 2571 | TGTTCTAAATTCATGCAATGTCCTAGAGACACTGGTGTGTTTAACAGCATTTGAAGAATG | 2630 |
| Qy | 2634 | CCAG-TACACCCTAATATGTTATTAGGATAGCATAGT-----AC | 2673 |
| Db | 2631 | CCAGTTACACCCTAATATGTTATTGAGGATAGGATAGTTGTGACAACTTGATATCCATAG | 2690 |
| Qy | 2674 | AGTGTAACTAGAGCAAGTGTATGCAATTTGATGCTTTATAGATAGTTATAGCATCTGGCAG | 2733 |
| Db | 2691 | AGGCAAGTCAATGTCAAATTTTGATGCTTTTATAGATAGTTTAAATAGGCATCTGGGAGA | 2750 |
| Qy | 2734 | AAGCTTGAATATTGATG-TCTAGAATGGTAAGACAACTCTGCGTGTCTAAAAAATAA | 2792 |
| Db | 2751 | AAGCTTGAATATTGATGTTCTAGAATGGTAAGACAACTCTGCGTGTCTAAAAAATAA | 2810 |
| Qy | 2793 | AAGGTAAAGACATCAAAATTTTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT | 2852 |
| Db | 2811 | AAGGTAAAGACATCAAAATTTTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT | 2870 |
| Qy | 2853 | GCGGAAGGGAACACAGTCACCGCTGATTTAAGAGTGAATTTTGCTGTGCGCATTTAGCCA | 2912 |
| Db | 2871 | GCGGAAGGGAACACAGTCACCGCTGATTTAAGAGTGAATTTTGCTGTGCGCATTTAGCCA | 2930 |
| Qy | 2913 | CTGTGATATGTTTAGGGGCTGCAGTGGCTGTCTAAACCTCCACTTGGGATTAAGGCTAAAG | 2972 |
| Db | 2931 | CTGTGATATGTTTAGGGGCTGCAGTGGCTGTCTAAACCTCCACTTGGGATTAAGGCTAAAG | 2990 |
| Qy | 2973 | AAGCTATGGACAAGGTAGTTTTTAAGAAACATATAGCAACAGAAATTTATAACAGCAGGGA | 3032 |
| Db | 2991 | AAGCTATGGACAAGGTAGTTTTTAAGAAACATATAGCAACAGAAATTTATAACAGCAGGGA | 3050 |
| Qy | 3033 | ATGGGTTCTTGATCTCTTTTGTTTCTTCTCTTATCTCTAGGAGAGCTGTCTTCTCATG | 3092 |
| Db | 3051 | ATGGGTTCTTGATCTCTTTTGTTTCTTCTCTTATCTTCTAGGAGAGCTGTCTTCTCATG | 3110 |
| Qy | 3093 | ACTTGGTGTGTGGGATATTATGATGAAGCCATGAAGAAAATTTCTATGCCAGAAAAGGTTTA | 3152 |
| Db | 3111 | ACTTGGTGTGTGGGATATTATGATGAAGCCATGAAGAAAATTTCTATGCCAGAAAAGGTTTA | 3170 |
| Qy | 3153 | TCCTTGATGTTTCCCTAGAACCTTGTTTCAAGCACAGAGGTGAGGTCCTTGGTCAATA | 3212 |
| Db | 3171 | TCCTTGATGTTTCCCTAGAACCTTGTTTCAAGCACAGAGGTGAGGTCCTTGGTCAATA | 3230 |
| Qy | 3213 | TGCACCGCTATATAAAGAGCTCTTTTTTGTATTATAGAGCTGTCTATATAAATGGACAGT | 3272 |
| Db | 3231 | TGCACCGCTATATAAAGAGCTCTTTTTTGTATTATAGAGCTGTCTATATAAATGGACAGT | 3290 |
| Qy | 3273 | TTCTATCATTTGATCACTTTTCTTACTAAAAATGGTGGCAGCTTGATGAATTTGTGGCCA | 3332 |
| Db | 3291 | TTCTATCATTTGATCACTTTTCTTACTAAAAAATGGTGGCAGCTTGATGAATTTGTGGCCA | 3350 |
| Qy | 3333 | AAACAAGTACTAAGATTGACAAGGTTCTAAATTTTGGCAATTTGATGATGCAATCTCGAAG | 3392 |
| Db | 3351 | AAACAAGTACTAAGATTGACAAGGTTCTAAATTTTGGCAATTTGATGATGCAATCTCGAAG | 3410 |
| Qy | 3393 | AACGAATTCACCGTCTGTTGGATCCACCCATCAAGTGGTAGATCTTTATCATACAAAATTTG | 3452 |
| Db | 3411 | AACGAATTCACCGTCTGTTGGATCCACCCATCAAGTGGTAGATCTTTATCATACAAAATTTG | 3470 |

| | | | |
|----|------|---|------|
| QY | 3453 | CTCCTCTAAGACTCTCTGGACCTGATGATGTAAGTCATACACAGATACTTGCCTCTCGCTT | 3512 |
| DB | 3471 | CTCCTCTAAGACTCTCTGGACCTGATGATGTAAGTCATACACAGATACTTGCCTCTCGCTT | 3530 |
| QY | 3513 | GCATTTGTCAGATACTCAGATTTTTTACCATTTTTTCAATATTTCTAAATAGATTTGGTACA | 3572 |
| DB | 3531 | GCATTTGTCAGATACTCAGATTTTTTACCATTTTTTCAATATTTCTAAATAGATTTGGTACA | 3588 |
| QY | 3573 | TATAAATGGTTGATGCTTGACGACATATGCGCCCTTTACCAATAATTTCCCTGCTCATCATCG | 3632 |
| DB | 3589 | TATATTTGTTTGATGCTTGACGACATATGCGCCCTTTACCAATAATTTCCCTGCTCATCATCG | 3648 |
| QY | 3633 | ATCATCGTAGTACTCTGCTTACTCTGTTTTTAAAGAACAAAAACATGAGCCATCAATCTTTTC | 3692 |
| DB | 3649 | ATCATCGTAGTACTCTGCTTACTCTGTTTTTAAAGAACAAAAACATGAGCCATCAATCTTTTC | 3708 |
| QY | 3693 | AAAAATAAAAAAGGTCATTGAGTACCTTCCTTGGCCACTGTTTAAATAGTTTGGTCTTCC | 3752 |
| DB | 3709 | -AAAAATAAAAAAGGTCATTGAGTACCTTCCTTGGCCACTGTTTAAATAGTTTGGTCTTCC | 3767 |
| QY | 3753 | CAAAATAGTTAAATAGGTAGTGTAATATGATATATATTTTTCTTGTTGGTTGTTCTCG | 3812 |
| DB | 3768 | CAAAATAGTTAAATAGGTAGTGTAATATGATATATATTTTTCTTGTTGGTTGTTCTCG | 3827 |
| QY | 3813 | TACCAAGAGTAAAAATGCACCTTTGTTGACAAAGAAATGATAGGCAGTCTTTGGCATGACAC | 3872 |
| DB | 3828 | TACCAAGAGTAAAAATGCACTTTGTTGACAAAGAAATGATAGGCAGTCTTTGGCATGACAC | 3887 |
| QY | 3873 | ACTTTTGCTTCCCTTTTCTGACAATATATGCTGTTTAAAGTGGCCATAAATAGATACATCGA | 3932 |
| DB | 3888 | ACTTTTGCTTCCCTTTTCTGACAATATATGCTGTTTAAAGTGGTCCATAAATAGATACATCGA | 3947 |
| QY | 3933 | CATGTTTTTTGTAGCAGGAGATTGATATTTGTTCTTATTTGCTTCCAAATTTAAAGCATATT | 3992 |
| DB | 3948 | CATG-TTTTTGTAGCAGGAGATTGATATTTGTTCTTATTTGCTTCC-ATTTAAAGCATATT | 4005 |
| QY | 3993 | CTTCTTTAGCAATGATTTCAATGTTGGGACATATTTGTGCTGCTATTTAAGTAAATTTGTTTG | 4052 |
| DB | 4006 | CTTCTTTAGCAATGATTTCAATGTTGGGACATATTTGTGCTGCTATTTAAGTAAATTTGTTTG | 4065 |
| QY | 4053 | ATATCATATATATCTTTAATTTGGTAATATATGTGCACATCTCGCTCCCTGATTTGCTTTTG | 4112 |
| DB | 4066 | ATATCATATATATCTTTAATTTGGTAATATATGTGCACATCTCGCTCCCTGATTTGCTTTTG | 4125 |
| QY | 4113 | TCTTTTCAAAAGGTTACTTGGAGACCCCTTAATTTCAAAGGAAGATGACACAGCTGCAGT | 4172 |
| DB | 4126 | TCTTTTCAAAAGGTTACTTGGAGACCCCTTAATTTCAAAGGAAGATGACACAGCTGCAGT | 4185 |
| QY | 4173 | ATTGAAGTCAAGGCTTGAAGCCCTTCCACGTCACAAACTAAGCCCTGTATGTTTCTTTAGCA | 4232 |
| DB | 4186 | ATTGAAGTCAAGGCTTGAAGCCCTTCCACGTCACAAACTAAGCCCTGTATGTTTCTTTAGCA | 4245 |
| QY | 4233 | ACTACGTTTTTAAATATTCAGATATCTTTTATAGGATGTAGTGCATCTCAGTTAAGCGGG | 4292 |
| DB | 4246 | ACTACGTTTTTAAATATTCAGATATCTTTTATAGGATGTAGTGCATCTCAGTTAAGCGGG | 4305 |
| QY | 4293 | ATTCCTTTCAAGTTGCATATACAGTGTTTTCTCTGTATATCTTTCAATGTTTTT | 4341 |
| DB | 4306 | ATTCCTTTCAAGTTGCATATACAGTGTTTTCTCTGTATATCTTTCAATGTTTTT | 4354 |

```

RESULT 2
; Sequence 2, Application US/09802927
; Patent No. US20010031236A1
; GENERAL INFORMATION:
; APPLICANT: UCHIMIYA, HIROFUMI
; APPLICANT: ARAI, SATOSHI
; APPLICANT: FUSHIMI, TAKAOMI
; APPLICANT: TAGAWA, MICHIO
; APPLICANT: FUKUZAWA, HIROMITSU
; TITLE OF INVENTION: DNA FRAGMENT HAVING PROMOTER FUNCTION

```

FILE REFERENCE: 204323USOCIP
; CURRENT APPLICATION NUMBER: US/09/802,927
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: PCT/JP98/04088
; PRIOR FILING DATE: 1988-09-10
; PRIOR APPLICATION NUMBER: PCT/JP99/04847
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Oryza sativa L.cv.Nipponbare
US-09-802-927-2

Query Match 31.9%; Score 1384.6; DB 10; Length 1443;
Best Local Similarity 99.2%; Pred. No. 1.9e-281;
Matches 1434; Conservative 0; Mismatches 4; Indels 7; Gaps 4;

QY 1 CTGCAGGAAGATTAAATAGCTGGACACACCAACCCCTGGTGGTGACGCCCTCTGTGT 60
DB 1 CTGCAGGAAGATTAAATAGCTGGACACACCAACCCCTGGTGGTGACGCCCTCTGTGT 60
QY 61 AATCAACTGGGGTTCGTTGGACATGGTTTTTTCAGGAAATTAAGCAAGAAAAATTAAG 120
DB 61 AATCAACTGGGGTTCGTTGGACATGGTTTTTTCAGGAAATTAAGCAAGAAAAATTAAG 120
QY 121 AAGATGCTCAAGCTGACATGAGAAACGTAATCCATGGAAGCGAATTCAGTCGTTTC 180
DB 121 AAGATGCTCAAGCTGACATGAGAAACGTAATCCATGGAAGCGAATTCAGTCGTTTC 180
QY 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTGCCTCAAC-GTTTGTAGTGGCTCCTATTG 239
DB 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTGCCTCAACGTTTGTAGTGGCTCCTATTG 240
QY 240 GCTCGTGATACGACTGTGTGTCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 299
DB 241 GCTCGTGATACGACTGTGTGTCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 300
QY 300 ACCGAAACCCCAAGAAATTTCTATTCCACCTCAGGGTAATTTGCTGCAACTATGCAAT 359
DB 301 ACCGAAACCCCAAGAAATTTCTATTCCACCTCAGGGTAATTTGCTGCAACTATGCAAT 360
QY 360 GAATACAAATTCGAAATATCATGTTATCTATCTGTCTCAAAATGAAATTTGAGTCA 419
DB 361 GAATACAAATTCGAAATATCATGTTATCTATCTGTCTCAAAATGAAATTTGAGTCA 420
QY 420 ACTGAGACTCAATACGATTTTCTTTTCAAAAGAAATTTATTAATTTTTCATGAA 479
DB 421 ACTGAGACTCAATACGATTTTCTTTTCAAAAGAAATTTATTAATTTTTCATGAA 480
QY 480 ACGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAAGTAGTCTAGTCGAGAAACAA 539
DB 481 ACGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAAGTAGTCTAGTCGAGAAACAA 540
QY 540 AATTAATATCACATAAAGAGAGTTGTTAATTAACCACTGTTTCTGCTACTCAACTC 599
DB 541 AATTAATATCACATAAAGAGAGTTGTTAATTAACCACTGTTTCTGCTACTCAACTC 600
QY 600 TAAATTTGTAATTTCTTATTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAAATAACG 659
DB 601 TAAATTTGTAATTTCTTATTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAAATAACG 660
QY 660 TAGACGCTAAGCCCAACCTATCAAGCTAAGTTTCAGAGGTGAAGTACGACGAGAAAT 719
DB 661 TAGACGCTAAGCCCAACCTATCAAGCTAAGTTTCAGAGGTGAAGTACGACGAGAAAT 720
QY 720 ATGATGTTTTTAAATATGATTTTTTTTAAATAACTTTTCACTAAATTTCTTTAGAAAC 779
DB 721 ATGATGTTTTTAAATATGATTTTTTTTAAATAACTTTTCACTAAATTTCTTTAGAAAC 780
QY 780 ATATCAATTAATGGTTGAAAAACGTCACATAGAAAACTAAGAACGATGAGTTGGAA 839
|||||

DB 781 ATATCAATTAATGGTTGAAAAACGTCACATA--GAAAACTAGAACGATGAGTTGGAA 838
QY 840 ACAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTGATTTTCAAA 899
DB 839 ACAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTGATTTTCAAA 898
QY 900 CGCATGATAACGAGAAAGCTCATTAGCAGCATTTACTTAGATATTATTAATTAATTAAC 959
DB 899 CGCATGATAACGAGAAAGCTCATTAGCAGCATTTACTTAGATATTATTAATTAATTAAC 958
QY 960 TTGAAAAAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATTTATTTTAAAAAA 1019
DB 959 TTGAAAAAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATTTATTTTAAAAAA 1018
QY 1020 CACACCAATTTAACCTTTTAAAGCATCTTAATAGGAACCGAGGAAGTTAAAGATTCAAC 1079
DB 1019 CACACCAATTTAACCTTTTAAAGCATCTTAATAGGAACCGAGGAAGTTAAAGATTCAAC 1078
QY 1080 CGAA---GTGCTTTGGATAATGAAAAATGGGTGGGATTAGAATTTGGTAAATCAATCAGGG 1136
DB 1079 CGAAGGTGTGTTGGATAATGAAAAATGGGTGGGATTAGAATTTGGTAAATCAATCAGGG 1138
QY 1137 TTAGGATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1196
DB 1139 TTAGGATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1198
QY 1197 TTTGGTGGGTAGAAAAATTTTCCCATACACTCCCGAGAGCGGTGCTGCTGGTGGG 1256
DB 1199 TTTGGTGGGTAGAAAAATTTTCCCATACACTCCCGAGAGCGGTGCTGCTGGTGGG 1258
QY 1257 CAGAAGCGTCTTTTTCGTTGGAAAAAAACTGCTTAAAAAGGAAACAGAAAGAGAGAGAG 1316
DB 1259 CAGAAGCGTCTTTTTCGTTGGAAAAAAACTGCTTAAAAAGGAAACAGAAAGAGAGAGAG 1318
QY 1317 CTTT-GTTGTCCCGTCTCACCAGAAACGAAACCAAAAGCCCAACCTAAACCTCTC 1375
DB 1319 CTTTGTGTTGTCACCGTCTCACCAGAAACGAAACCAAAAGCCCAACCTAAACCTCTC 1378
QY 1376 GATCCGACCGAGACTCTCTCCATTTTCAGCGCGCAGCGGAGAGACGCGACGCGAGTTGCG 1435
DB 1379 GATCCGACCGAGACTCTCTCCATTTTCAGCGCGCAGCGGAGAGACGCGACGCGAGTTGCG 1438
QY 1436 TCGAC 1440
DB 1439 TCGAC 1443
|||||

RESULT 3

US-09-878-574-3892
; Sequence 3892, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plance
; FILE REFERENCE: 38-21(15401) B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3892
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-004-Q1-B1-E12
US-09-878-574-3892

Query Match 2.8%; Score 122; DB 10; Length 404;
Best Local Similarity 80.3%; Pred. No. 2e-16; Indels 0; Gaps 0;
Matches 143; Conservative 0; Mismatches 35;

;; PRIOR FILING DATE: 2000-01-27
;; NUMBER OF SEQ ID NOS: 999
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 484
;; LENGTH: 894
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
;; NAME/KEY: misc feature
;; FEATURE:
;; LOCATION: (1)...(894)
;; OTHER INFORMATION: n = A,T,C or G

US-09-770-445-484

Query Match 2.3%; Score 101.8; DB 10; Length 894;
Best Local Similarity 71.9%; Pred. No. 4.9e-12;
Matches 133; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 3303 AAATGGTCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAAGGTTCTCAA 3362

Db 583 AAGCAGAGAAGCTTGTATGAGATGCTTAATAGAAGGGAGCTCAGATAGATAAGGTGCTTA 524

Qy 3363 ATTTTGCATTCATCAATACATCTGGAAGACGAATTACCGTCTGTTGGATCCACCAT 3422

Db 523 ATTTTGCATTCATCAATTCGGTCTCGAAGAAAGAAATTACTGGAAGGTGGATTACCCCTT 464

Qy 3423 CAAGTGGTAGATCTTATCATACAAAATTGCTCTCTTAAGACTCTCTGACTTGATGATG 3482

Db 463 CAAGTGGAGAGCTATCATACTAAATTTCGCACCTCTTAAGTTCCAGGAGTCGATGATC 404

Qy 3483 TAAAT 3487

Db 403 TGAAT 399

RESULT 7

US-09-923-876-3999
; Sequence 3999, Application US/09923876
; Patent No. US20020013958A1

;; GENERAL INFORMATION:
;; APPLICANT: Lalgudi, Raghunath V.
;; APPLICANT: Kamigaki, Laura Y. (Ito)
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;; FILE REFERENCE: PL-0012-1 CON
;; CURRENT APPLICATION NUMBER: US/09/923,876
;; CURRENT FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: 09/298,329
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/085,331
;; PRIOR FILING DATE: 1998-05-05
;; NUMBER OF SEQ ID NOS: 6332
;; SOFTWARE: PERL Program
;; SEQ ID NO 3999
;; LENGTH: 292
;; TYPE: DNA
;; ORGANISM: Zea mays
;; NAME/KEY: misc feature
;; FEATURE:
;; OTHER INFORMATION: Incyte ID No. US20020013958A1 700454231H1

;; LOCATION: 34, 59, 83, 88, 92, 113, 122, 142, 145, 158, 172-173, 217, 236-237, 255,
;; LOCATION: 260, 277-278, 284
;; OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-3999

Query Match 2.1%; Score 92.8; DB 10; Length 292;
Best Local Similarity 73.2%; Pred. No. 2.5e-10;
Matches 123; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

Qy 2838 TAGGTCCACCTGGCTGCGGAAGGGAACACAGTCACCGCTGATTGAAGGATGAATTTGCT 2897

Db 109 TCGNCCACCTGGTCTCTGGAAGGGTACTCAGTNCCTTATTGAAGGAAATATTGCC 168

Qy 2898 TGTCACATTTAGCCACTGGTGATATGTTAGGGCTGTCAGTGGCTGCTAAAACTCCACTTG 2957
Db 169 TGTNNCAITTTAGCCACTGGTGATATGTTAGGGCTGCTGTCGACCCANGAC-AACTCTAG 227
Qy 2958 GGATTAAGGCTAAAGAAGCTATGGACAAGGTAGTTTTTTAAGAAACATA 3005
Db 228 GTATCAAAANNTAAAGAAGCTATGGATANGGANAGCTTTGTTTCAGATA 275

RESULT 8

US-09-923-876-6281
; Sequence 6281, Application US/09923876
; Patent No. US20020013958A1

;; GENERAL INFORMATION:
;; APPLICANT: Lalgudi, Raghunath V.
;; APPLICANT: Kamigaki, Laura Y. (Ito)
;; APPLICANT: Sherman, Bradley K.
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;; FILE REFERENCE: PL-0012-1 CON
;; CURRENT APPLICATION NUMBER: US/09/923,876
;; CURRENT FILING DATE: 2001-08-06
;; PRIOR APPLICATION NUMBER: 09/298,329
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/085,331
;; PRIOR FILING DATE: 1998-05-05
;; NUMBER OF SEQ ID NOS: 6332
;; SOFTWARE: PERL Program
;; SEQ ID NO 6281
;; LENGTH: 267
;; TYPE: DNA
;; ORGANISM: Zea mays
;; NAME/KEY: misc feature
;; OTHER INFORMATION: Incyte ID No. US20020013958A1 700458807H1

;; NAME/KEY: unsure

;; LOCATION: 47, 77, 98, 174
;; OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-6281

Query Match 2.1%; Score 92; DB 10; Length 267;
Best Local Similarity 81.9%; Pred. No. 3.5e-10;
Matches 104; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1466 GCTGCGGAGATGGCGGGAACCTGGAGGACGTGCGGTGCGAGCTGATGACGAGCTG 1525

Db 42 GCGGTNGAGATGGCGGGAATCTGGAGGACGTGCANTCACTGGATCTGATGACGAGCTG 101

Qy 1526 CTCGCGCCCATGAAGTGCAGCTCCAGCCGACACAGCCGCTCATCTCGTGGTAACGCC 1585

Db 102 CTCGCGCCCATGAAGTGCAGCTCCAGCCGACACAGCCGCTCATCTCATACGGTCCACC 161

Qy 1586 CGCTCT 1592

Db 162 TGGCTCT 168

RESULT 9

US-09-887-576-828/c
; Sequence 828, Application US/09887576
; Patent No. US2002014047A1

;; GENERAL INFORMATION:
;; APPLICANT: Budworth, P.
;; APPLICANT: Brown, D.
;; APPLICANT: Chang, H.
;; APPLICANT: Zhu, T.
;; APPLICANT: Han, B.
;; APPLICANT: Wang, X.
;; APPLICANT: Cooper, Bret
;; TITLE OF INVENTION: Promoters for regulation of plant expression
;; FILE REFERENCE: 1360.001US1
;; CURRENT APPLICATION NUMBER: US/09/887,576
;; CURRENT FILING DATE: 2001-06-25
;; PRIOR APPLICATION NUMBER: US 60/213,848

```
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 828
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-887-576-828

Query Match      1.4%; Score 62.4; DB 10; Length 2000;
Best Local Similarity 61.9%; Pred. No. 0.0013;
Matches 99; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY  903 ATGATACAGCAAGCTCATTAGCACATTATTACTAGATATTTATTAATAAATCTG 962
Db  1227 ACGTAAACAGGATGAGCCATTACGTTATGATTAATTAAGTATTAACCTATTTTAAACTTT 1168
QY  963 AAAAAATATTATTGTAATTTTAAACAATGTATGCATAAAATTTTAAAAAACAC 1022
Db  1167 AAAATAGATTAATATGATTTTAAAGCAACTTCTATAGAAATTTTGGAAAAAC 1108
QY  1023 ACCAATTTAACCCCTTTAAAAAGCATCTTAATAGGAACGA 1062
Db  1107 ACACCGTTTAGTGTCCGAAACGTGCGCGTGGAAATGA 1068

RESULT 10
US-09-815-242-6911
; Sequence 6911, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6911
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
; US-09-815-242-6911

Query Match      1.4%; Score 57; DB 10; Length 639;
Best Local Similarity 59.6%; Pred. No. 0.011;
Matches 96; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY  2834 ATTTTAGGTCCACCTGGCTGGGAAAGGAAACACAGTCACCGCTGATTAAGATGAATTT 2893
Db  13 ATTTAGGTCCACCTGGCTGGGAAAGGAACTCAAGCAAAAATCGTAGAACAAATTC 72
QY  2894 TGCCTTGTCGCAATTTAGCCACTGGTGATATGTTAGGGGCTGCAGTGGCTGCTAAAACTCCA 2953
Db  73 CATGTTGCACATATCTCAACAGGTGATATGTTCCGCGTGCATGGCAAACTCAACTGAA 132
QY  2954 CTTGGGATTAAGGCTAAAGAGCTATGGACAAGGTAGTGT 2994

Query Match      1.4%; Score 61.6; DB 10; Length 645;
Best Local Similarity 59.9%; Pred. No. 0.0012;
Matches 106; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY  2835 TTTTAGGTCCACCTGGCTGGGAAAGGAAACACAGTCACCGCTGATTAAGGATGAATTT 2894
Db  14 TTTTAGGTCCACCTGGCTGGGAAAGGAACTCAAGCAAAATTTATTAAGCAAAATTTG 73
QY  2895 GCTTGTGCCATTTAGCCACTCGTGATATGTTGAGGGCTGCAGTGGCTGCTAAAACTCCAC 2954
Db  74 GTATCCCGCAAAATTTCAACTCGTGATATGTTCCGTGCTGCAATCAAAGCGGGGACTGAAC 133
QY  2955 TTGGGATTAAGCTAAAGAACTATGGACAAGGTAGTGTGTTTAAAGAACATATAGCAACAG 3014
Db  134 TTGGCAAAACAGCTAAAGCATTAATGGATGAAGGTAAATTAGTGCCAGATGAATTAACCG 193

RESULT 11
US-09-815-242-9121
; Sequence 9121, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9121
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(639)
; US-09-815-242-9121

Query Match      1.3%; Score 57; DB 10; Length 639;
Best Local Similarity 59.6%; Pred. No. 0.011;
Matches 96; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY  2834 ATTTTAGGTCCACCTGGCTGGGAAAGGAAACACAGTCACCGCTGATTAAGATGAATTT 2893
Db  13 ATTTAGGTCCACCTGGCTGGGAAAGGAACTCAAGCAAAAATCGTAGAACAAATTC 72
QY  2894 TGCCTTGTCGCAATTTAGCCACTGGTGATATGTTAGGGGCTGCAGTGGCTGCTAAAACTCCA 2953
Db  73 CATGTTGCACATATCTCAACAGGTGATATGTTCCGCGTGCATGGCAAACTCAACTGAA 132
QY  2954 CTTGGGATTAAGGCTAAAGAGCTATGGACAAGGTAGTGT 2994
```

```

Db      133 ATGGGTGTTCTTGAAGTCATATTGAAGGGTGAAAT 173
        ||| || | |||| | ||| |||||| | |||
RESULT 12
US-08-910-366A-4
; Sequence 13, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 213
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte ID No. US20020137081A1 415620CBI
US-10-044-090-213
Query Match          1.3%; Score 57; DB 12; Length 908;
Best Local Similarity 62.1%; Pred. No. 0.013;
Matches 90; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY   2838 TAGGTCACCTGCCTCGGGAAGGCAACAGTCACCCTGATAAGGATGAATTTTCT 2897
DB   114  TGSGGCCCTCCCGGGCGCGTAAGAAGGACCCAGCACCATTTCTGTG 173
QY   2898 TGTGCCATTTAGCACTGCTGTATGTTGAGGGCTGCAGTGGCTCAAACCTCCACTTG 2957
DB   174  TGTGCCATTTAGCTACTGCGGACATCTGAGGCCCCATGTTGCTCTGCTCAGAGTAG 233
QY   2958 GGATTAAGGCTAAAGAACGCTATGGA 2982
DB   234  GAANAAGCTGAAGGCNACTATGGA 258
RESULT 13
US-08-910-386A-4/c
; Sequence 4, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
```

```

; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME: 11
; CHROMOSOME/SEGMENT: 11q, RG103
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2367..4205
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member D"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 4201..9071
; OTHER INFORMATION: /note= "retrofit, a copia-like,
; OTHER INFORMATION: transposon-like element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4484..8821
; OTHER INFORMATION: /product= "retrofit"
; OTHER INFORMATION: /gene= "gag/pol"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 9915..11712
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 10020..10975
; OTHER INFORMATION: /note= "Krispie, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 12626..12750
; OTHER INFORMATION: /note= "Pop-012, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice2, transposon-like
; OTHER INFORMATION: element"
;
; US-08-910-386A-1
;
; Query Match 1.3%; Score 55.8; DB 8; Length 13341;
; Best Local Similarity 57.0%; Pred. No. 0.065;
; Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
;
; QY 891 ATTTTCAACGATGATTAACGAGAGCTCATTAGACATTTACTTAGATATTATA 950
; Db 1842 AATTGACAAGATCGTGAACGAAATAGCTGTAGTGATATAATTAATTAAGTAGTAATT 1783
; QY 951 ATTATAAATTTGAAAAAATATTATTATTTGAATTTTTTAAACAATGTATGCATAAAATTATT 1010
; Db 1782 ATTTTAAACTTTAAAAATAGATTAATATATTTTAAAGAAACTTCTCTATAGAAAGTTT 1723
; QY 1011 TTTTAAAAACACACCAATTTTAAACCCCTTTAAAAAGCATCTCTAATAGGAAACGAGGAAGTT 1069
; Db 1722 TTTCAAAAAACACACCACTTTAAACAGTTTCTGAAAAACATACCGTGAAAAACGATGCTTTT 1664
;
; RESULT 15
; US-08-910-386A-6/c
; Sequence 6, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME: 11
; CHROMOSOME/SEGMENT: 11q, RG103
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5213..18201
; OTHER INFORMATION: /note= "Xa21 gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(5213..7889, 8732..9132)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 disease resistance gene"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 9645..9769
; OTHER INFORMATION: /note= "Pop-Oil1, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice1, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(15118..17720, 17827..18204)
```

; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "x21 gene family member C; 2 bp
; OTHER INFORMATION: deletion causing frame-shift mutation of
; OTHER INFORMATION: ORF compared to family member A1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16183..16184
; OTHER INFORMATION: /note= "location of 2 bp deletion
; OTHER INFORMATION: compared to family member A1"
; US-08-910-386A-6

Query Match 1.3%; Score 55.8; DB 8; Length 19639;
Best Local Similarity 57.0%; Pred. No. 0.076;
Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
Qy 891 ATTTTCAACGCATGATAAAGCAGAGAAAGCTCATTAGCACATTATTACTTAGATATTATA 950
Db 4688 AATTGACAGATACGTAGAACGAAATAGCTGTTAGTGTATAATTAATTAGTAGTAATT 4629
Qy 951 ATTATAAACTTGAAAAAATATTTATTTGAATTTTTTAAACAATCTATGCATAAAATTATT 1010
Db 4628 ATTTTAACTTAAAAAATAGATTATATATTTTAAAGAACTTTTCTATAGAAAGTTT 4569
Qy 1011 TTTTAAAAACACCAATTTAACCCCTTTAAAGAGCATCCTAATAGAAACGAGGAGTT 1069
Db 4568 TTTCAAAAAACACACCATTTTAAACAGTTCTGAAACATACGCGTGAAAAACGATGCTTTT 4510

Search completed: March 15, 2003, 20:31:01
Job time : 469.605 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:16:40 ; Search time 3558.28 Seconds
(without alignments)
11777.621 Million cell updates/sec

Title: US-09-802-937-2
Perfect score: 1440
Sequence: 1 ctgcagagattaattagg.....cgcgacgagtgctgtcgac 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| | |
|-----|-----------------|
| 1: | gb_ba.* |
| 2: | gb_htg.* |
| 3: | gb_in.* |
| 4: | gb_om.* |
| 5: | gb_ov.* |
| 6: | gb_pat.* |
| 7: | gb_ph.* |
| 8: | gb_pl.* |
| 9: | gb_pr.* |
| 10: | gb_ro.* |
| 11: | gb_sts.* |
| 12: | gb_sy.* |
| 13: | gb_un.* |
| 14: | gb_vi.* |
| 15: | em_ba.* |
| 16: | em_fun.* |
| 17: | em_hum.* |
| 18: | em_in.* |
| 19: | em_mu.* |
| 20: | em_om.* |
| 21: | em_or.* |
| 22: | em_ov.* |
| 23: | em_pat.* |
| 24: | em_ph.* |
| 25: | em_pl.* |
| 26: | em_ro.* |
| 27: | em_ste.* |
| 28: | em_un.* |
| 29: | em_vi.* |
| 30: | em_htg_hum.* |
| 31: | em_htg_inv.* |
| 32: | em_htg_other.* |
| 33: | em_htg_mus.* |
| 34: | em_htg_pin.* |
| 35: | em_htg_rod.* |
| 36: | em_htg_mam.* |
| 37: | em_htg_vrt.* |
| 38: | em_sy.* |
| 39: | em_htgo_hum.* |
| 40: | em_htgo_mus.* |
| 41: | em_htgo_other.* |

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-----------|--------------------|
| 2 | 1373.6 | 95.4 | 98782 | 2 | CNS08CB5 | AL731756 Oryza sat |
| 3 | 115.6 | 8.0 | 137348 | 8 | AP003297 | AP003297 Oryza sat |
| 4 | 109.2 | 7.6 | 124815 | 2 | AP005594 | AP005594 Oryza sat |
| 5 | 109.2 | 7.6 | 138025 | 8 | AP003261 | AP003261 Oryza sat |
| 6 | 109.2 | 7.6 | 169663 | 2 | CNS07YFX | AL713952 Oryza sat |
| 7 | 108.8 | 7.6 | 126659 | 8 | OSJN00015 | AL606455 Oryza sat |
| 8 | 108.8 | 7.6 | 155582 | 2 | OSJN00094 | AL606649 Oryza sat |
| 9 | 108.4 | 7.5 | 100419 | 2 | OSIG00047 | AL732355 Oryza sat |
| 10 | 107.6 | 7.5 | 100635 | 8 | AP003446 | AP003446 Oryza sat |
| 11 | 107.6 | 7.5 | 144074 | 8 | AP003302 | AP003302 Oryza sat |
| 12 | 106.4 | 7.4 | 156054 | 8 | AB023482 | AB023482 Oryza sat |
| 13 | 104.8 | 7.3 | 129838 | 8 | AC079038 | AC079038 Oryza sat |
| 14 | 103.6 | 7.2 | 170025 | 2 | OSJN00031 | AC079038 Oryza sat |
| 15 | 102 | 7.1 | 131983 | 8 | AC023240 | AC023240 Oryza sat |
| 16 | 101 | 7.0 | 143959 | 2 | AC090055 | AC090055 Oryza sat |
| 17 | 100 | 6.9 | 6735 | 8 | RIC3H3M | L28995 Oryza sativ |
| 18 | 100 | 6.9 | 129845 | 2 | AP004165 | AP004165 Oryza sat |
| 19 | 99.4 | 6.9 | 121615 | 2 | AP003981 | AP003981 Oryza sat |
| 20 | 99.4 | 6.9 | 146670 | 8 | AP003250 | AP003250 Oryza sat |
| 21 | 99 | 6.9 | 142596 | 8 | AP003314 | AP003314 Oryza sat |
| 22 | 99 | 6.9 | 166753 | 8 | AP002866 | AP002866 Oryza sat |
| 23 | 98.4 | 6.8 | 159893 | 8 | OSJN00095 | AP002866 Oryza sat |
| 24 | 98.4 | 6.8 | 189834 | 8 | OSJN00171 | AL606652 Oryza sat |
| 25 | 96.8 | 6.7 | 108805 | 2 | CNS08CBF | AL662974 Oryza sat |
| 26 | 96.8 | 6.7 | 150238 | 2 | CNS08CA9 | AL731875 Oryza sat |
| 27 | 96.6 | 6.7 | 151100 | 8 | AP003453 | AL831796 Oryza sat |
| 28 | 96.2 | 6.7 | 136553 | 2 | AC093953 | AP003453 Oryza sat |
| 29 | 95.8 | 6.7 | 132407 | 8 | OSJN00117 | AL607002 Oryza sat |
| 30 | 95.2 | 6.6 | 139566 | 2 | AP004161 | AL606652 Oryza sat |
| 31 | 95.2 | 6.6 | 147954 | 2 | AP003983 | AP004161 Oryza sat |
| 32 | 95.2 | 6.6 | 165394 | 8 | AC025296 | AP003983 Oryza sat |
| 33 | 94.8 | 6.6 | 160174 | 2 | AP005103 | AC025296 Oryza sat |
| 34 | 94.8 | 6.6 | 162010 | 2 | AP004121 | AP005103 Oryza sat |
| 35 | 94.6 | 6.6 | 102119 | 2 | AP003919 | AP004121 Oryza sat |
| 36 | 93.2 | 6.5 | 122559 | 2 | AP004301 | AP003919 Oryza sat |
| 37 | 92.8 | 6.4 | 168698 | 2 | OSJN00023 | AP004301 Oryza sat |
| 38 | 92.4 | 6.4 | 148985 | 8 | AP002094 | AL606588 Oryza sat |
| 39 | 92.2 | 6.4 | 135874 | 8 | AC113337 | AP002094 Oryza sat |
| 40 | 92 | 6.4 | 142776 | 2 | AP005109 | AC113337 Genomic s |
| 41 | 92 | 6.4 | 167110 | 2 | AC120505 | AP005109 Oryza sat |
| 42 | 91.4 | 6.3 | 1154 | 8 | RICADKA | AC120505 Oryza sat |
| 43 | 91.4 | 6.3 | 133595 | 2 | AC129227 | D10334 Oryza sativ |
| 44 | 91.4 | 6.3 | 164679 | 8 | AC078944 | AC129227 Oryza sat |
| 45 | 91 | 6.3 | 137462 | 8 | AP002538 | AC078944 Genomic S |

ALIGNMENTS

| | | | | | | |
|------------|------------|--|---------|-----|--------|-----------------|
| RESULT 1 | AB041773 | AB041773 | 4341 bp | DNA | linear | PLN 20-APR-2000 |
| LOCUS | AB041773 | Oryza sativa Adk-a gene for adenylylate kinase, partial cds. | | | | |
| DEFINITION | AB041773 | Oryza sativa Adk-a gene for adenylylate kinase, partial cds. | | | | |
| ACCESSION | AB041773.1 | GI:7630192 | | | | |
| VERSION | AB041773.1 | GI:7630192 | | | | |
| KEYWORDS | | adenylate kinase. | | | | |
| SOURCE | | Oryza sativa DNA. | | | | |
| ORGANISM | | Oryza sativa | | | | |
| | | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | |
| | | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; | | | | |
| REFERENCE | | 1 (bases 1 to 4341) | | | | |
| AUTHORS | | Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S. | | | | |
| TITLE | | Oryza sativa Adk-a gene, promoter region and partial ORF | | | | |

Weissenbach, J. and Quetier, F.
Oryza sativa chromosome 12 sequencing
Unpublished
REFERENCE 2 (bases 1 to 98782)
Genoscope.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
On Jul 11, 2002 this sequence version replaced gi:20372820.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* by the accession number will be preserved.
FEATURES
source
1..98782
Location/Qualifiers
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OJ1220 D01"
/clone_lib="Monsanto"
BASE COUNT 30065 a 19192 c 20278 g 29238 t 9 others
ORIGIN
Query Match 95.4%; Score 1373.6; DB 2; Length 98782;
Best Local Similarity 99.2%; Pred. No. 6.2e-234;
Matches 1434; Conservative 0; Mismatches 4; Indels 8; Gaps 5;
Qy 1 CTCGAGAGAGATTAAATTAGTGGACACACCAAAACCTGTGTGGTGAGCCCTGTGTGT 60
Db 63789 CTCGAGAGAGATTAAATTAGTGGACACACCAAAACCTGTGTGGTGAGCCCTGTGTGT 63848
Qy 61 AATCAACTGGGTGTCGTTGGACATGTTTTTCAGAGAAAATTAAGCAAGAAAATTAAAG 120
Db 63849 AATCAACTGGGTGTCGTTGGACATGTTTTTCAGAGAAAATTAAGCAAGAAAATTAAAG 63908
Qy 121 AAGAAATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTTTC 180
Db 63909 AAGAAATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTTTC 63968
Qy 181 TCTTGACTACCACTGTTTAAAGATAATAGACAGTGGCCAAC - GTTTGATGGTTCCTATTG 239
Db 63969 TCTTGACTACCACTGTTTAAAGATAATAGACAGTGGCCAACGGTTTGAITGGCTCCTATTG 64028
Qy 240 GCTCGTGTGATACTGACTTGTGTCAACAAGCATCAAAATTCCTCTTGGAGTATCTTTATT 299
Db 64029 GCTCGTGTGATACTGACTTGTGTCAACAAGCATCAAAATTCCTCTTGGAGTATCTTTATT 64088
Qy 300 ACCGAAAAACCCAAAGATTATTCTATTCCACCTCAGGGTAATTTGTGCTGAACATATGCAAT 359
Db 64089 ACCGAAAAACCCAAAGATTATTCTATTCCACCTCAGGGTAATTTGTGCTGAACATATGCAAT 64148
Qy 360 GAATACAAATTCGCAAAATATCATGGTTATCTATCTTGTCTCAAAATTCGAAATTTGAGTCCA 419
Db 64149 GAATACAAATTCGCAAAATATCATGGTTATCTATCTTGTCTCAAAATTCGAAATTTGAGTCCA 64208
Qy 420 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTTTCATGAA 479
Db 64209 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTTTCATGAA 64268
Qy 480 ACCGAAATCAACCGTTTCGAGAAATATGCTGTATATAATAAGTAGTACTAGTCGAGAAACAA 539
Db 64269 ACCGAAATCAACCGTTTCGAGAAATATGCTGTATATAATAAGTAGTACTAGTCGAGAAACAA 64328

| | | | |
|-----------|--|------|--|
| KEYWORDS | | | |
| SOURCE | Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0698A10. | CDS | join(6553..6815,6898..7498) /gene="P0698A10.3" /note="contains ESTs C73695 (E20178), AU082323 (E20178)" /codon_start=1 /product="putative myb-related protein" /protein_id="BAB92433.1" /db_xref="GI:20804748" |
| ORGANISM | Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza. | gene | complement(join(9381..9448,10362..10453,10723..10803,11573..11701,12144..12277,13341..13677,16691..16813,16960..17116,17804..17874,17967..18004)) /gene="P0698A10.4" /codon_start=1 /translation="MGRSPPCEKKAHTNKAWTKBEDORLIAYIKAHGEGCWRSUPKAA GLRCGSKRLRWNYLRPDILKRGNTDDDDDELIIKHALGNKWSLIAGQLPQRTDN EIKYWNTHIKRLLSRGIDQTHRPVSAGSSAAAAGSLTTASTAAPPSLAPAPPQ QHRLNHPVHAAPGNASGAPSPGSDGHSSGSSGDAPRCPLNDLDLDSMSL PSSPKTPAAASSTTASRHHHQQTICLCYHLGVRNGDVCSCAAAPSPAGPAPRA FLRLSEGOVI" |
| REFERENCE | Sasaki, T., Matsumoto, T. and Yamamoto, K. | | |
| AUTHORS | Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0698A10 | gene | join(19424..19431,19465..19629,20505..20545,20966..21079,21836..21852) /gene="P0698A10.5" /db_xref="GI:20804749" |
| TITLE | Published Only in Database (2001) | CDS | join(19424..19431,19465..19629,20505..20545,20966..21079,21836..21852) /gene="P0698A10.5" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92433.1" /db_xref="GI:20804750" |
| REFERENCE | Sasaki, T., Matsumoto, T. and Yamamoto, K. | | |
| AUTHORS | Submitted (19-FEB-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan | CDS | join(23333..24646,25086..25210,25302..25443,25546..25806) /gene="P0698A10.6" /db_xref="GI:20804751" |
| TITLE | (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) | gene | complement(join(202..450,580..654)) /gene="P0698A10.1" /db_xref="GI:20804746" |
| COMMENT | On May 15, 2002 this sequence version replaced gi:16416397. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0698A10 clone has an overlap with P0648C09 (DDBJ: AP003922) at the position 1 to 92,190 of 5' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/genomeseq.html. | CDS | join(23333..24646,25086..25210,25302..25443,25546..25806) /gene="P0698A10.6" /db_xref="GI:20804751" |
| FEATURES | Location/Qualifiers | | |
| source | 1. .137348 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="1" /clone="P0698A10" complement(join(202..450,580..654)) /gene="P0698A10.1" complement(join(202..450,580..654)) /gene="P0698A10.1" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92433.1" /db_xref="GI:20804746" | gene | complement(join(5591..5603,5744..5886)) /gene="P0698A10.2" /db_xref="GI:20804747" |
| Gene | translation="MRYPRVQCAPVNRSSMDSGRIHLFVGEAFARSVVAGWYGRPR AARAPRRGRPLNKGAAQLAPRSYARDVRVEISLICVTLLRRRRRRPPYAPTAATR RGRTP" | CDS | join(5591..5603,5744..5886) /gene="P0698A10.2" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92432.1" /db_xref="GI:20804747" |
| gene | translation="MEKGEEREGRSGEGEAGHAGALIGAVDSFRPVEAGDSNVRG WRWIGR" | gene | join(27824..27925,28000..28127,28637..28780,28932..29064) /gene="P0698A10.7" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92437.1" |
| CDS | join(6553..6815,6898..7498) /gene="P0698A10.3" | CDS | join(27824..27925,28000..28127,28637..28780,28932..29064) /gene="P0698A10.7" /note="hypothetical protein" /codon_start=1 /protein_id="BAB92437.1" |

```
/db_xref="GI:20804752"
/translation="MSGSPPAAPAAAEIALASWANKLERDRVATRRWEGPAIAKPARTI
WGPPIRRSTAVATVTRTPDSPSRPRGHGRGATRRMLRLSLVLPANACRIIA
VNHARRIGNQKTRTTPRCCHAIRVWFMRLRAAAAAAASVLRQFTWMPDLCNVQRAY
EQIQCYYV"
complement(join(31579..31693,31909..32131,32235..32469,
32585..32716,32809..33174))
/gene="P0698A10.8"
CDS
complement(join(31579..31693,31909..32131,32235..32469,
32585..32716,32809..33174))
/gene="P0698A10.8"
/codon_start=1
/product="putative DNA polymerase I"
/protein_id="BAB92438.1"
/db_xref="GI:20804753"
/translation="WPPPAVSTAPVAPVAVSSSSAAAAAPRRPSRLRVTPYASVSS
SSSPSTFAVISIPPPPAHRSKQLVRDGAAPAKPRVFLDVNPLCFRGSORSIGA
FARWALFAHVSRLDPVAVLDGEGNEYRRLPLPSYKAHRPGPGTGADSRVIDVL
RCNPFVVRVDGAEADDVVATLQVLRGFRVIGSPDKQFKQLISDVQVMPPE
IGRWFSYLRHVYAQKCDPTADLSLRCFIGDEADGPIQHLVPGFRKTAVALLLK
HGSLENLNTAAVRTVGKDYAQDALVKHADYLRKNYEVLRLKRDVKVQLDRLWLSTRD
SCNDSVLSDFLTKEGONIR"
complement(34177..34410)
/gene="P0698A10.9"
CDS
complement(34177..34410)
/gene="P0698A10.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92439.1"
/db_xref="GI:20804754"
/translation="WTSUTGRAFDNRDNGFIPTATELARSMADTNTDLHVTSPTPAARR
SWPPFSLLTASRRMPGKQKQKTRKRLKEI"
complement(join(35312..35473,36485..36646))
/gene="P0698A10.10"
CDS
complement(join(35312..35473,36485..36646))
/gene="P0698A10.10"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92440.1"
/db_xref="GI:20804755"
/translation="MEHGRGREGRETGWVVDAPADTCESGYVWAGPFFWAVST
NRLRLSDILQESGIPVPIQFLCQOLASKRLNFSENRSTRDRFPKTPSDSKYSND
ETURL"
join(37075..38148,39504..39906,40314..40415,40513..40736)
/gene="P0698A10.11"
CDS
join(37075..38148,39504..39906,40314..40415,40513..40736)
/gene="P0698A10.11"
/note="contains ESTs AU101336 (E11055), AU101337 (E11055)"
/codon_start=1

Query Match      8.0%; Score 115.6; DB 8; Length 137348;
Best Local Similarity 76.3%; Pred. No. 2.6e-11;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 893 TTTCAAACGCATGATAACGAGAAAGCTCATTAGCACTATTACTTAGATATTATAAT 952
|||||
Db 136817 TTTTAACTCAGCAAAACGAGAAACTTATTAGCATAATTAATTAAGTATTATAAT 136758
|||||

QY 953 TATAAATCGAAAAATATTTATTGAATTTTAAACAATGTATGCATAATTTT 1012
|||||
Db 136757 TATAAATCGAAAAATATTTATTGAATTTTAAATAAATTTTATATAGATTTT 136698
|||||

QY 1013 TTAATAACACACCAATTTAAACCCCTTTAAAGCATCTCTAATAGAAACGAGGAAGTAAA 1072
|||||
Db 136697 TTAATAACATATCATTTAAGTTTGAAGGATCTACTAACGGAACGATAAATAATAA 136638
|||||

QY 1073 GATTCA 1078
|||||
Db 136637 GTTTGA 136632

RESULT 4
AP005594

LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 9 clone P0701E06,
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION
AP005594
VERSION
AP005594.1 GI:21952963
KEYWORDS
HTG; HTGS PHASE2
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0701E06.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 Sasaki,T., Matsumoto,T., Hattori,M., Sakaki,Y. and Katayose,Y.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
clone:P0701E06
JOURNAL
Published Only in Database (2002)
REFERENCE
2 (bases 1 to 124815)
AUTHORS
Sasaki,T., Matsumoto,T., Hattori,M., Sakaki,Y. and Katayose,Y.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
tel:81-298-38-7441, fax:81-298-38-7468)
COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source
1. .124815
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="P0701E06"
BASE COUNT 37055 a 25766 c 25390 g 36503 t 101 others
ORIGIN

Query Match      7.6%; Score 109.2; DB 2; Length 124815;
Best Local Similarity 63.2%; Pred. No. 3.5e-10;
Matches 168; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 820 TAAGAACGATGAGTTGGGAAACAAGAGAAAAACACAGCCCTTAAAGCTTCTTGATCCTCTA 879
|||||
Db 56107 TTATAAGTATCTGTAGGATTAAATCAATTAAAGATGACTAATGCTCGGTCAATCACCTC 56166
|||||

QY 880 GTTGGAGTTGATTTTCAACGCATGATAACGAGAAAGCTCATTAGCACATATTACTT 939
|||||
Db 56167 ATTTGGAGATGAGATTTTCAACGCACACAAACGAGAAATCTCATTAGCACATGATCAAT 56226
|||||

QY 940 AGATATTTATATTAATAAATTTGAAAAAATATTTATTGAAATTTTAAACAATGTATG 999
|||||
Db 56227 AAGTGTTAATTAATTAATAAATTTGAAAAATGGAATTTATTGTATATTATAACTACTTCTA 56286
|||||

QY 1000 CATAAATTTATTTTAAACACACCAATTTTAACCCCTTTAAAAAGCATCTCTTAATAGGAAA 1059
|||||
Db 56287 TATAGATTTTATTTTAAAGAAACACATCATTTAACTGTTTGAAGAAAGATGCTTAATGGAAGA 56346
|||||

QY 1060 CGAGGAAGTTAAAGATTCCACGAAGT 1085
|||||
Db 56347 TCAGAGAAAGTTGAAGTTTACTAGAGT 56372
|||||

RESULT 5
AP003261/c
LOCUS
```

DEFINITION *Oryza sativa* (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0471B04.

ACCESSION AP0033261

VERSION AP003261.3 GI:22202658

SOURCE *Oryza sativa* (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0471B04.

ORGANISM *Oryza sativa* (japonica cultivar-group)

SpERMatoPhyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; *Oryza*.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE *Oryza sativa nipponbare*(GA3) genomic DNA, chromosome 1, PAC clone:P0471B04

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 138025)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

COMMENT (E-mail:tsasakienias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)

On Aug 9, 2002 this sequence version replaced gi:21202837.

Genes were predicted from the integrated results of the following: GENSCAN1.0, BLAST2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, nr.

(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0471B04 clone has an overlap with P0698A10 clone (DBJ : AP003297) at the position 1 to 24,438 of 5' end and an overlap with P018C10 clone (DBJ: AP003227) at the position 48,801 to 138,025 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/Genomeseq.html>.

Location/Qualifiers

1. 138025

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="1"

/clones="P0471B04"

join(9533. .9715,9887. .10019,10850. .10988,11365. .11401, 13640. .13931,14823. .14955,15564. .15576)

/gene="P0471B04.1"

join(9533. 9715,9887. .10019,10850. .10988,11365. .11401, 13640. .13931,14823. .14955,15564. .15576)

/gene="P0471B04.1"

/note="hypothetical protein

similar to Arabidopsis thaliana chromosome 1, F10C21_14"

/codon_start=1

/protein_id="BAC07317.1"

/db_xref="GI:22202659"

/translation="MSPPSVIGQFDGDTTYTKVFGGLAWETQKTMKRYFEQGEILLE AVITDKNTGRSGYGVTRFDPAAMRACVDPAVIDGRRANCLASLGVQSRPPT PQHCARSFRVMKSFQOAGIQGLGAAPSHAFPHYAI PQGLPYHYGVSPYSPDY SPYTNYN1YXGAOY PFGVAAAAAAGAWTGSFPFYFQFGQSGSTTTNVASQO GYNLYQPMHFSTVASTAAVTGAQYGGPLSLAASPAQAQAKANGKSNPNQAL TQGGFMUQANWTIKVTSKLIQECHFKISEKP"

complement (join(17199. .17342,17914. .18008,18141. .18178,

18784. .19073))

/gene="P0471B04.2"

complement (join(17199. .17342,17914. .18008,18141. .18178, 18784. .19073))

/gene="P0471B04.2"

/codon_start=1

/product="putative SnRK1-interacting protein 1"

/protein_id="BAC07318.1"

/db_xref="GI:22202660"

/translation="MASGALARLLGRRAAATPLLARPPFAAKASRRPOEPAPFSD EDDPAGGEVAAPAPTEGISKPLAEVLRELGVKRYPEALVKTRVEDGFSKLKYPWHIVNK LNIHAPGEVSIRVSISSDGKSVSVIYRVTLHGTDAEIYREATGTSSADDDTGYGDP VQAEAWAFRRACRTRIGLGLHLHEDMS"

join(24593. .24925,25029. .25403)

/gene="P0471B04.3"

join(24593. .24925,25029. .25403)

/gene="P0471B04.3"

/note="hypothetical protein"

/codon_start=1

/protein_id="BAC07319.1"

/db_xref="GI:22202661"

/translation="MARRGGGVILRGRCRRRGERGEQVRPLPLASTVTGGRGGR GRPRRERGACGTARQSSSPRSSGRRSSSSMTSRSEQHLLSTDOSRCLSNWMS TGLPVSSKVIRORGAGVRRRGRSGRRGRRGAGAAVQDGGWRGCGHGRGRSRGA GMAANDGADAAVAADADGLGGCGRSFPIRRRAALLWQSSSSRRRFPF1PTLPAAS PSAVLLSSGAMLPNRR"

join(25836. .26033,26372. .26471,26764. .26924,27200. .27397, 27476. .27568,27766. .27888,28671. .28918,29038. .29209, 29441. .29527,29557. .29820,30449. .30611,30873. .30960, 31044. .31113,31600. .31728,32356. .32445)

/gene="P0471B04.4"

join(25836. .26033,26372. .26471,26764. .26924,27200. .27397, 27476. .27568,27766. .27888,28671. .28918,29038. .29209, 29441. .29527,29557. .29820,30449. .30611,30873. .30960, 31044. .31113,31600. .31728,32356. .32445)

/gene="P0471B04.4"

/note="hypothetical protein

similar to Arabidopsis thaliana chromosome 5, At5g49880"

/codon_start=1

/protein_id="BAC07320.1"

/db_xref="GI:22202662"

/translation="MILRTPQKRRAADSDVDAANAATTRSPVSDRLVLYDRPTAL VPATQPEGMDDMVCTYHCRQVKSFPVALDTAEKQVQYRATLDDMEERLKSDEE RACQDKLNVEQEELAAKRGSAQMERLKEKVEDQFOERYCDQIKKIGETLUKKEI DSRIAAESSASAKESVKEGLNQLSENSEKKAALKKLSYLQDDTKLSIKLNA ELBRMLRALNSDEAKLLNEQLDKQJDESVRNEMEHRLNLLSSSYERTPSD DQKLLLOEELANYKEKVEDEARLLKSSHTNVLLKEKILEEQCSRAEMLSKLOE IEAKQKLELEASCTALLSNIPDVSSFGDIPQKIADLOKALTNLKNKVEVTSQJKE LKVALEFADLUSKQRAEATLAKERASATREIKRLELLAAISSEDRDLRKDAVSK SRGDDDASSKFCAPILNGMHNISCEVTIQNMESDLSRMKVKVTELESTIRDORELIS QHTEUNLMNEKLSIESRKAKSLERDQDLRSOVALLESGLHGDSASSTKVLRMVN TLAVDNEAKOTIEALQAEKTKERLQAEVQELQADVTVDVNIAEKLAQLKNQIA TLEKREERYKVAFAERISVFRKACCSLFGYKIVMNDQOQSGIPVTRFILHSVVAQSD DEKLEFDYESGNTNIVVDIFIRKMNSIPATANTLMTSEFNKRSIC"

join(33470. .33475,33573. .33794,33969. .34033,34129. .34322, 34573. .34839,34942. .35029,35125. .35250,35325. .35434, 35557. .35705)

/gene="P0471B04.5"

join(33470. .33475,33573. .33794,33969. .34033,34129. .34322, 34573. .34839,34942. .35029,35125. .35250,35325. .35434, 35557. .35705)

/gene="P0471B04.5"

/note="contains ESTs C91865 (E31978), AU166151 (E31978)"

/codon_start=1

/product="putative elicitor response protein"

/protein_id="BAC07321.1"

/db_xref="GI:22202663"

/translation="MMTLTSERQPEQKPPRTPPLPKGKAAALCVASFVGLLLS GRVVVPLLPGGSSPASNYKTSFTGCENKRAKIGESNPTDINNEVSRTHHAISLDK AVSLEMELAEVERASAAVAGTAVSGPKAFVVGINTAFSSKRRDSLRDTWV PRGDKRLRLKEKGIIVRFVIGSGAAAGDGLDRAVDAEDAKENGLFLRDJHVEGYH ELSKTRVYITAVATWDADFYVYKDDVHVNLGMLTSRLAKYTRPRVYVCKMKGSP VLSQKGVYHEPBYWKGDEGNKFRHATCQITAYVSKDLAAYISINQILHRFANEDV

FEATURES

source

gene

CDS

gene

```
SIGAWLIGLEVEHVDNRSMCCATPPDCWKKRAGNVCVASFDMSCGVCYKSDVRMKHI
HRACGEGQAVWSVAT"
join(36438..36570,36637..36956)
/gene="P0471B04.6"
CDS
join(36438..36570,36637..36956)
/gene="P0471B04.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC07322.1"
/db_xref="GI:2202664"
/translation="MSLLTPSLVVKPSPRCPRQYRGRGSMSSVCGPGEEHHWG
PAERREYRRRRGRRRVLPVGNNGACPLSPRCGADGATPRSRKGGPDAP
TAPLPDPVTVRGEVSDGQVQWQWQDQGREREREATAD"
complement(join(38223..38304,39187..39287))
/gene="P0471B04.7"
complement(join(38223..38304,39187..39287))
/gene="P0471B04.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC07323.1"
/db_xref="GI:2202665"
/translation="WVKNDRSVTGLTPVTCWDTNNAPTSHLLPAEWSTFHSIDFT
CKVIKRIIVKYSATIG"
gene
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/gene="P0471B04.8"
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/gene="P0471B04.8"
/note="contains ESTs AU057294 (S21318), AU057295 (S21318)"
/codon_start=1
/product="putative male sterility 1 protein"
/protein_id="BAC07324.1"
/db_xref="GI:2202666"
/translation="WVNSRPLKRAITRVEARDFAGFPAGDGGAGTFRVAVRGFLA
RYALRPLVPSFPAAAPPHLLTRVSLRVGDEEGGGGAVELNVVEEDVLR
SYVQDQRPVGSFPCGRYHFIENDNNQVCGKRSCCLRCGCTPTVAGESRCLC
NFDMPGELEECGYMHLDDNTHLLHAVHANGYHLRVNGEGGRSLCTGRDMSFW
DLCKVLHVQVTVMDISKHGMVYRLHAIITSGHPWYGEWYKFGAGSALTSDTQV
EAVDTLSGIALYFSPHQIRTPLONTIALYALSRLQVTRDLPFRFIMHLLHQR
KNETSKEPTDHEKSVANLVCKWKEDIIDRAETAMKLVQVQPGQWVRALRGAR
SKAVSDQLDYSRLGKGLMDDGHFIIVRNCVSAIEYRLSDNSQSVDAAFGP
SYDHLHLKFLYNALNPETMLASQPEVIGASSHAAAKLIDCKQFKHYDQHTPRA
PLNPLFLLSRVCSIELLDHPKDYTAPELVLLPASATIAELKIOTRAFQTYLMFOS
YQVEOLPDPNPSIDTLVKHVLGSSQLVRGRCTGDNRRIVQFRMERGLENTWVDC
CGAKDDGGERMLACDVGWQHTRCSGISDFDVPKEFICRKCASPRGKRGGGG
GSRMDVSAAGRCKDEIGSSVGGAGKFRMATVG"
join(54852..54959,55057..55188,55283..55412,55519..55605,
56032..56102)
/gene="P0471B04.9"
join(54852..54959,55057..55188,55283..55412,55519..55605,
56032..56102)
/gene="P0471B04.9"
Query Match 7.6%; Score 109.2; DB 8; Length 138025;
Best Local Similarity 76.3%; Pred. No. 3.5e-10;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 893 TTTCACCGCATGATAACGAGAACTCATAGACATATTACTAGATATTATAT 952
Db 23907 TTTTAACTCACCAAAACGAGAAAACCTTATTAGCATAATTAATTAAGTATTAAT 23848
QY 953 TATAAACTTGAAAAAATATTTATTTGAATTTTTTAAACAATGATGCAATAATTTT 1012
Db 23847 TATAAACTTGAAAAAATATTTATTTGATTTTTTAAATAAATCTTTATATAGATTTT 23788
QY 1013 TTAATAACACACCAATTTAACCCCTTTAAAAAGCATCCTTAATAGGAACGAGGAAGTTAA 1072
Db 23787 TTAATAACATATCAITTTAACAGTTTGAAGAGCATACTACGAGAAACGATTAATTA 23728
QY 1073 GATTCA 1078
Db 23727 GTTTGA 23722
```

```
RESULT 6
CNS07YPX
LOCUS
DEFINITION
CNS07YPX 169663 bp DNA linear HTG 27-JUL-2002
Oryza sativa chromosome 12 clone OJ1618_C05, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION
AL713952
VERSION
AL713952.2 GI:22003292
KEYWORDS
HTG; HTGS PHASE2; HTGS_ACTIVEFIN.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 169663)
Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Segrens,B., Pellatier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 169663)
Genoscope.
Direct Submission
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jul 29, 2002 this sequence version replaced gi:19715885.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1.169663
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OJ1618_C05"
/clone_lib="Monsanto"
BASE COUNT 46256 a 37378 c 36384 g 49645 t
ORIGIN
Query Match 7.6%; Score 109.2; DB 2; Length 169663;
Best Local Similarity 68.8%; Pred. No. 3.3e-10;
Matches 150; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 861 AAGCTCTTTGATCCTCTAGTTCGAGGCTTCTTCAACGCATGATAACGAGAAAGCT 920
Db 103185 AAGCCTAGGTTTCGGCTTGTTCGATAGAAAATTTTAAACGCGCAAAACAAGAGCT 103244
QY 921 CATTAGCACATTTACTCTAGATATTTATATTAATAACTTGAAAAAATATTTATTGTA 980
Db 103245 CATTAGCACATAATTGATTAGTATTATTAATATATAAATTTGAAAAATGATTTATTGTA 103304
QY 981 ATTTTAAACAATGATGATGATTAATTTTAAAAAACAACCAATTTAAACCCCTTAA 1040
Db 103305 TTTTAAACAACCTCTATATAGAAAATTTTCTTAAAAAACAACATCATTTAAACAGTTAA 103364
QY 1041 AAAGCATCTTAATAGGAACGAGAAAGTTAAAGATTCA 1078
Db 103365 AAAGCGTGCTAACAGAAACGAGGAAGTTGAAGTTTGA 103402
```

```

RESULT 7
OSJN00015          126659 bp   DNA       linear      PIN 17-JUL-2002
LOCUS              Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB00079B02,
complete sequence.
ACCESSION          AL606455
VERSION            AL606455.2   GI:21740404
KEYWORDS           HTG
SOURCE             Oryza sativa.
ORGANISM           Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
AUTHORS            Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X.,
Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L.,
Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C.,
Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B.,
Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G.,
Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F.,
Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y.,
Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., and
Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and
Hong, G.F.
TITLE              Direct Submission
JOURNAL            Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
REMARK             Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB00079B02.
COMMENT            On Jul 12, 2002 this sequence version replaced gi:15552693.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM
(http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
complete sequence against NCBI none redundant protein database (nr)
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES
source
1..126659
/organism="Oryza sativa"
/variety="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OSJNB00079B02"
/clone_lib="CUGI-OSJNB"
BASE COUNT        35469 a 27283 c 28042 g 35865 t
ORIGIN
Query Match       7.6%; Score 108.8; DB 8; Length 126659;
Best Local Similarity 55.9%; Pred. NO. 4.1e-10;
Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;
QY 660 TAGACGGCTAACCCACCACCATTAAGCTTAAGTTCGAGAGGTGAAGTACGCACGCAAAAT 719
DB 15094 TAGATTGAAACCTTTTCCTTCCTAGCACGTAAACGAAACAACATATTAGCCGCAATGA 15153
QY 720 ATGATGGTTTATTATATGATTTTTTTTAAATAAATTTTACATAAAATTTCTTTAGGAAC 779
DB 15154 TTAGCAATATATTATATATATTATTGAAGCACTTCTTTTATTATTTCATAAAGAC 15213
QY 780 ATATCATTTAATGGTTTGAACACGTGCAATAGAAACTAGAACGATGAGTTGGGAA 839
DB 15214 GCATCATTTTAGTCGTTTGGAAACGTGCACGCTGAAACAAATAGACTAGAGTTAAAAA 15273
QY 840 ACAAGAGAAAAACACACGCTTAAGGCTTCTTGATCTCTCTAGTTGGAGGTTGATTTTCAA 899

```

```

Db 15274 ATTGGAAGAAATATACTAAGGCCCGTTCGATCTCCAGAGTGGAGATAGATTTTAT 15333
QY 900 CGCATGATAAAGCAAGAGCTCATTAGCACATTTACTTAGATATTATTATTAATTAAC 959
Db 15334 GGCAGAAAACGAGAAATCTCATTAACACATGATTAATTAATTAATTAATTAAT 15393
QY 960 TTGAAAAAATATTATTATGAAATTTTTTAAACAA--TGTATGCATAAATATTATTTTAAA 1017
Db 15394 TTGAAATATGATTTATTATGCTTTTTTAAACAACACTTGTTATATAGAACTTGCTGTA 15453
QY 1018 AACACACCAATTTAAACCTTTTAAAGACATCCTTAATAGGAACGAG 1063
Db 15454 TAATTTACCTTTTAACCTTTTGAAGAAACGGTCTGATCGAAATAAG 15499

RESULT 8
OSJN00094/c
LOCUS              Oryza sativa chromosome 4 clone OSJNB0003B01, *** SEQUENCING IN
DEFINITION         PROGRESS ***, 3 ordered pieces.
ACCESSION          AL606649
VERSION            AL606649.2   GI:21912479
KEYWORDS           HTG; HTGS PHASE2.
SOURCE             Oryza sativa.
ORGANISM           Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
AUTHORS            Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X.,
Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L.,
Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C.,
Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B.,
Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G.,
Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F.,
Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y.,
Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J.,
Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and
Hong, G.F.
TITLE              Direct Submission
JOURNAL            Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
REMARK             Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB0003B01.
COMMENT            On Jul 19, 2002 this sequence version replaced gi:15594106.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a PHASE2 sequence. Gaps are shown by nnnn. Genes
were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM
(http://genemark.biology.gatech.edu/GeneMark/), tRNAscan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
complete sequence against NCBI none redundant protein database (nr)
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 61924: contig of 61924 bp in length
* 61925 62124: gap of 200 bp
* 62125 76571: contig of 14447 bp in length
* 76572 76771: gap of 200 bp

```



```

* 76772 155582: contig of 78811 bp in length.
FEATURES             Location/Qualifiers
  source             1..155582
                    /organism="Oryza sativa"
                    /variety="Nipponbare"
                    /sub_species="japonica"
                    /db_xref="taxon:4530"
                    /chromosomes="4"
                    /clone="OSJNBh0003B01"
                    /clone_lib="CUGI-OSJNBb"
BASE COUNT          40185 a 36681 c 38162 g 40148 t 406 others
ORIGIN
Query Match          7.6%; Score 108.8; DB 2; Length 155582;
Best Local Similarity 55.9%; Pred. No. 4e-10;
Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

Qy 660 TAGACGGCTAAGCCACCCATCAAGCTAAGTTCGAGAGGTGAAGTACGACGAAATAAT 719
Db 14837 TAGATTGAGAACCTTTCTCTCTAGCACGTAAAAACGAAACACATATTAGCGCATAAGTA 14778

Qy 720 ATGATGGTTTAAATATGATTTTTTTTAAATAACTTTCACATAAATTTCTTTAGGAAC 779
Db 14777 TTAGCAATATATTAAATAAATTTTGAAGCAACTTCTTTTAAATTTTATTTCATAAAGAC 14718

Qy 780 ATATCATTTAATGGTTTGA AAAACGTGCACATAGAAAACCTAAGAACGATGAGTTGGGAA 839
Db 14717 GCATCATTTTGTGCGTTGGAGACGTGCAGTGA AAAACAATAGACTAGAGTTAAAAAA 14658

Qy 840 ACAAGAGAAAAACACAGCCTTAAGGCTTCTTGATCTCTAGTTGGAGGTGATTTTCAA 899
Db 14657 ATTGGAAGAAAATAACTAAGGCCCGCTCGATCTCCAGGAGTGAGATATTTTAT 14598

Qy 900 CGCATGATTAACGAGAAAGCTATTAGCACATTTATTCTAGATATTTTATAATTATAC 959
Db 14597 GGCACGAAAACGAGAAATCTCATTAACCATGATTAAATTAATAATTAATTTAAT 14538

Qy 960 TTGAAAAAATATTATTATTTGAATTTTTTAAACAA--TGTATGCATAAATATTTTTTAA 1017
Db 14537 TTGAAAAAATGATTATTATTTCTTTTTTAAACAACTTGATATAGAACTTGTTGTA 14478

Qy 1018 AACACCAATTTAAACCTTTAAAGCATCTTAATAGGAAACGAG 1063
Db 14477 TAATTTACCTTTTAACTGTTTGA AAAACGTGCTGATGGAATAAAG 14432

RESULT 9
OSIG00047          100419 bp DNA linear HTG 01-AUG-2002
LOCUS              Oryza sativa chromosome 4 clone H0322F07, *** SEQUENCING IN
DEFINITION         PROGRESS ***, in ordered pieces.
ACCESSION          AL732355
VERSION            AL732355.1 GI:22091860
KEYWORDS           HTG; HTGS PHASE2.
SOURCE             Oryza sativa
ORGANISM           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS            Han, B., Chen, Z.H., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q.,
                  Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J.,
                  Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y.,
                  Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K.,
                  Zhou, B., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P. and
                  Hong, G.F.
Direct Submission
Submitted (05-MAY-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
REMARK             Oryza sativa indica (guangluai4) genomic DNA, chromosome 4, BAC
                  clone: H0322F07.

```

```

Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
This is a complete sequence.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES             Location/Qualifiers
  source             1..100419
                    /organism="Oryza sativa"
                    /variety="guangluai4"
                    /sub_species="indica"
                    /db_xref="taxon:4530"
                    /chromosomes="4"
                    /clone="H0322F07"
                    /clone_lib="BAC-HindIII"
BASE COUNT          29642 a 21538 c 21324 g 27915 t
ORIGIN
Query Match          7.5%; Score 108.4; DB 2; Length 100419;
Best Local Similarity 72.9%; Pred. No. 5.1e-10;
Matches 153; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 869 TTGATCTCTAGTTGGAGTTGATTTTCAAAACGCATGATAAAGCAGAAAGCTCATTAGCA 928
Db 51985 TTGATCATCGTTTGGCAGATAGATTTTGTCTACGCAAAACGAGAAATCTCATTAGCA 52044

Qy 929 CATTTACTTTAGATATTTTATAATTAATAACTTTGAAAAAATATTATTATTGAATTTTT 988
Db 52045 CATGATTAATTAAGTATTAAGTATTATAAAATTTTAAAAATGAATTTATTG-CTTTTT 52103

Qy 989 AACATGTATGCATAAATTTTAAAAAACACACCAATTTAACCTTTAAAAAGCATC 1048
Db 52104 AACACTTGTATTAATAACTTTTTTAAAAAACGTACCATTTAACGTTTGA AAAAGCGT 52163

Qy 1049 CTAATAGAAACGAGAGTGTAAAGATTCA 1078
Db 52164 CTAATGAAAAACGAGAAAGTTGAAGTTGA 52193

RESULT 10
AP003446           100635 bp DNA linear PLN 23-MAY-2002
LOCUS              Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION         BAC clone: OJ1529_G03.
ACCESSION          AP003446
VERSION            AP003446.3 GI:21104777
KEYWORDS           Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
SOURCE             clone: OJ1529_G03.
ORGANISM           Oryza sativa (japonica cultivar-group)
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS            Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE              Oryza sativa nipponbare (CA3) genomic DNA, chromosome 1, BAC
                  clone: OJ1529_G03
JOURNAL            Published Only in Database (2001)
REFERENCE          2 (bases 1 to 100635)
AUTHORS            Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE              Direct Submission
JOURNAL            Submitted (28-MAR-2001) Takuji Sasaki, National Institute of
                  Agrobiological Sciences, Rice Genome Research Program; Kannondai
                  2-1-2, Tsukuba, Ibaraki 305-8602, Japan
                  (e-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
                  Tel:81-298-38-7441, Fax:81-298-38-7468)
                  On May 22, 2002 this sequence version replaced gi:16041055.
                  Genes were predicted from the integrated results of the following:
                  GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
                  (October 1998 version). The genomic sequence was searched against

```

NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21m13 to 13mrev of the BAC
clone. This sequence of OJ1529 G03 clone has an overlap with
P0454H12 clone (DDBJ : AP003255) at the position 1 to 31,196 of 5'
end and an overlap with P0703B11 clone (DDBJ : AP003302) at the
position 85,137 to 100,635 of 3' end. Detailed information on
overlap and assembly quality together with annotation of this entry
is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

FEATURES

source

```

1. 100635
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OJ1529 G03"
join(28..136,803..1007,2468..2542,2784..3209,3357..3502,
4325..4443)
/gene="OJ1529_G03.1"
join(28..136,803..1007,2468..2542,2784..3209,3357..3502,
4325..4443)
/gene="OJ1529_G03.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB93365.1"
/db_xref="GI:21104778"
/translation="MLRPVDNVHTEKRAQKLDCTDEYISLSVMWIFFSPERRTRAH
FALGRKLIQVNMNLLKKEKPEPLAGAKTIIQTTAEGSLNKYLGNKPEPLKKE
GEDQTVRHKSQVHKVODFELLVVVVRKNNKELTLETNPEOTNRTILLESKKRTC
NOTQTDPLQSEQRFAAPRSGSTVRRGARPEIALTNELGKRERAPHLHPQRPQR
RGYTYTNAATFQVQLSHSNPIRTPLPFLDRGSYDEGEVEAKLWRTSED
RVFRTIGRAARKGSGEESRAGKESAAVQPSATAFACSGRTESVFDTIQVKNKRAD
FSNNLWDFTKNTYDFIPNNLGS"
complement(join(5387..5424,5515..5711,6719..7027,
7922..8039,8200..8246,9341..9463,11128..11291))
/gene="OJ1529_G03.2"
complement(join(5387..5424,5515..5711,6719..7027,
7922..8039,8200..8246,9341..9463,11128..11291))
/gene="OJ1529_G03.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB93366.1"
/db_xref="GI:21104779"
/translation="MPAFAPDETDNRNKPPELDITLLLLPSLPVPEAARRRGWE
SPAARDPRREYVHQHRIWLDNRMAACPRLNRSGVRGARIRIARTARDKHGYK
YSTKADWAPLKLHILVKHYCVINCEVIGLECSWIKVSSLVESRPRPLSRV
RSRRLAGALSPSLAAGCDYGLDLHLAASNDLGNRRRAAABOEDDLKRLVRV
HAGMGQCNLPFSFSSFLNCFVFLIETHRAILERKIIAYLTTRKGEAKDRCPK
STYTKGVCTFYRRCENALICSOPTVLFDFPKVTKHSNLMLVAESQVKYGANL"
join(13718..14497,14615..15031)
/gene="OJ1529_G03.3"
join(13718..14497,14615..15031)
/gene="OJ1529_G03.3"
/note="hypothetical protein"
similar to Arabidopsis thaliana chromosome 3, F17014.14"
/codon_start=1
/protein_id="BAB93367.1"
/db_xref="GI:21104780"
/translation="MSTSTIAKTTSQAQNSRPSPTPTSPRIMSSSSSTSAAGRPS
SGRTSTTSRSTSPVPSRSSRSTPTQPIRTSSAPLAGHSVGRIFGNNI
TIGRPVTSNGRSPASSAPSSRPSNSRARAPVRPLDIPDPSPTPNLRKLPOR
PLSAGRARPVGGLGPKAPNAEQVRSAPVKMTVPAITRSKFPDAPSRVSLTNGHOS

```

```

RQSGSAVEGQHTKPLRSITGTDNGFGORTISRKSLDMAIKHMDIRQNLGGIRGASLFFP
PSIRSAKVRSAKSPAHLSGDCHTFDNGSINGHFGSDNGSTLNNNGSGSTGSP
DREISGHTKELSELDTVTARYEALLREDITNTNMLHNVDDKSDQSQSFVDFHRPEPLP
EPFGPL"
join(21779..23066,23415..24040)
/gene="OJ1529_G03.4"
join(21779..23066,23415..24040)
/gene="OJ1529_G03.4"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="BAB93368.1"
/db_xref="GI:21104781"
/translation="MQDHITAFVLVLSLFCAPPAKSAADNSDKQALLAFAASLPHG
KLNWSAAPVCTSWGVTCTPDNSRVQTLRPAGVGLFGPLSPDTLGLDALEVLRLR
SRNTWDPPEVGSIPSLHSYLOHNNLSGIPTSLTSLTFLDLSYNTFGEIPLRV
ONLTQLTALLQNNLSGPIPDQLPKLHLNLSNNLSGPIPSLORFPANSFLGNA
LLLCFIPQPCGTAPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LLLCFIPRKRKSTTASSSKGTAGGKNGKPEEYSSGQAEKNNKLVFFEGCSY
NFDELDLLRASAELVSGSYGTYTVALEDDGTTVVVKRLKEVVVGKDFEQOQVTVGR
VGHQNVVPLRAYYSKDEKLLVYDIPSGSLAVVLHGNKATGAPLDMETRVKISLG
VARGIATHAEGGKFTIHNKSSNILLSONLDCGVSEFGLAQLMTIPAPARLVGYR
APEVLETKPTOKSDVYSGVLVLEMLTGKAPLSPGREDSEIHLPRVQSVVREMT
AEVEDVLLRHPNIEDEMQVLVAMACVAAAPPQPKMDEVIRIVEIRINSYSGRS
PPEEKQDESAAAP"
complement(join(24926..25054,25703..25828,25943..26270,
26491..26580,26648..26724,26883..26957,27041..27142,
27226..27344,28604..28727,28791..28841,28954..29052,
29468..29578))
/gene="OJ1529_G03.5"
complement(join(24926..25054,25703..25828,25943..26270,
26491..26580,26648..26724,26883..26957,27041..27142,
27226..27344,28604..28727,28791..28841,28954..29052,
29468..29578))
/gene="OJ1529_G03.5"
/note="hypothetical protein"
similar to Arabidopsis thaliana chromosome 1, T20H2.13"
/codon_start=1
/protein_id="BAB93369.1"
/db_xref="GI:21104782"
/translation="MGIVSMVLSFGICGFSVIGVYFLYFQPTDVKNVKVRPL
VEYDSDSLDTEIPLMKNPDYDLMDLNRLEFLMWPYLKALICTAODANPPIA
KNEKYKIDIKETFTLGLSPPTFQVQVQVNEQELVMEPSLKWAGNPNTTVVKA
YGLKATIQVDMQVFLPRITLPLVSSFFCFANILVSLMEKPHVDLGLLGDADMA
IPVLVKFQVETIMDVASFMFLPKTLEVPDPSKASKPVGILLVFLVRAQNLREKQ
PLGRDPFIPLYQFGKDEKLSKLKLTPTGVTITDNLKTWEPNGIQNEKSA
GETITLTYKPFKEGNIQKEDPGGLLYVVAEAKELGKCNTPYVKLTFKGVKTKT
VVKENRNPKRSEFECECTPANDKLHVEVLGTKKALIRNKESLGHIDISLADVIN
KRIEMVDLNSRQGIQIEFQWKS"
complement(join(30702..30830,31071..31202,31304..31654,
31755..31983,32072..32148,33005..33079,33163..33264,
33366..33484,34116..34239,34340..34390,34507..34605,
35115..35225))
/gene="OJ1529_G03.6"
complement(join(30702..30830,31071..31202,31304..31654,
31755..31983,32072..32148,33005..33079,33163..33264,
33366..33484,34116..34239,34340..34390,34507..34605,
35115..35225))
/gene="OJ1529_G03.6"
/note="hypothetical protein"
similar to Arabidopsis thaliana chromosome 1, At1g20080"
/codon_start=1
/protein_id="BAB93370.1"
/db_xref="GI:21104783"
/translation="MGIVSMIDPDSGFCGFSAGIVGYFLYFQTEVDRKVCPL
VEYDNSLDGILHEIPLWKNPDIDRVDNRFLMWPYLKALICTAODANPPIA
ENKEKYKIDSIEPETLTLSLPTTFQMKAYVTEEOELIMEPSLKWANPNVTUVKA
YGLKATIQVVDLQVFPASPRITLPLVATIFPCAKILVSLMEKPHVDLGLLGDADMA
IPILYSFQVETIKKQVSMYLMWPKTLEVPDPSKASKRPVGLLVLVRAQNLQKDK
LKGSDPYVLKLELPEETKVTYTNLLKNDPNPENWEDKFVVTDPTELEIKVGGHK
KMGNNKILLKLEPPEETKVTYTNLLKNDPNPENWEDKFVVTDPTELEIKVGGHK
DVQGTDVGAAPGTAGGGLLYVVAEADLEGKHTNPVARIIFKGNKTKVVIKK
NRDPRVEEETEFVCEDDPANDKLHVEVLSPKPKKLLHGKTLGYIDINLAKVSKR
INITYHLIDSKTQIQIELQWRTS"

```



```
/protein_id="BAB85285.1"
/db_xref="GI:18844815"
/translation="MEPRKGAIGLLARPRVAHLGRVYTIQIWIISDVAMHTSGWAKI
KRKESPTLOQSNESHRYACKTARRGADGVGRKAMARTRLEEGNPGCGDEGRS
GAARHPQPPQEGGGLVAGSYSVFVLSLQH".22934)
/gene="P0703B11.5"
/join(18955)..19061,19338..19445,22352..22934)
/gene="P0703B11.5"
/join(18955)..19061,19338..19445,22352..22934)
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB85286.1"
/db_xref="GI:18844816"
/translation="MVKRSDNMDSSSECSGAKRRLQLQDSRYQENAMKKVICGTRT
EYVAPYHDGQWYKYGOKMIRGNSFPYCYVRCYVHODHGPCASKVHQNSDEPPLF
RVIYTNBHTCCTNSADYMASNIOIADASLRKAQAERLKAQVETPRLMHSP
PRCSGYNMANKERDVIYLSLLTVIRGCHTAESGNGNSAAALPNRPPPAVARSDHY
SCVAISPELLUPASDDVDFMLDSVLDPHWVEFLDLAWLKSEHTHG"
/join(23961..24007,24256..24442,25834..26617,27725..27794,
29077..29191)
/gene="P0703B11.6"
/join(23961..24007,24256..24442,25834..26617,27725..27794,
29077..29191)
/notes="contains EST D15357(C0514)
unknown protein"
/codon_start=1
/protein_id="BAB85287.1"
/db_xref="GI:18844817"
/translation="MWSGFPDNDNDSIRKICRIRVRSYSILIRDSPLYFPYS
SPFYPTPKMKNMTRALSIRSFIRKEGLHCSTAIKTSAGVNVHCTGAPAG
YPAATSPGASDITRCAAITHRPSSTPAKTRPQOLPAATPSRHRRLRRGLP
AEPTSPANRHOLLRTPTSPAHVATSSAASCPATVTVSGASHPVSPASAP
HRRQGLPPSPHATGSSGQIOISQSRSTPATPGATVRROQKTPPPATYLPDPC
ADLVYTSVTWTSKRCHLQTSRPHRTTPMPTSPPLDPAGRHGSRPRRERVA
PLPRALSSNDFLAGQLSWLTANVHKRDESIFDLAGEHESEEDYPVVVYSEIQ
TAISTTVR"
/join(29568..29636,30926..31120,31516..31684,32551..32689,
33006..33072)
/gene="P0703B11.7"
/join(29568..29636,30926..31120,31516..31684,32551..32689,
33006..33072)
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB85288.1"
/db_xref="GI:18844818"
/translation="WQCOROLNPKSPYACDASAVIOTATARIIRBGDGGHMPARE
CGGLTIKREGNSGDEDSWEARAKMVDERTVPGGSLTTTGVKTSWFTGVGDGN
VLDVATLEAFRRTIPKPISSDFCLSCSGASDCHMPLADPREGEADACYLAAAV
TETSKLGDDDDVQAALLQIRCFPPATGARGAAVFPVVGKSGTFVGFPTGY"
/join(34279..34442,39279..39383,40143..40713)
/gene="P0703B11.8"
/join(34279..34442,39279..39383,40143..40713)
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB85289.1"
/db_xref="GI:18844819"
/translation="WASPLKREOSFQFEASAOEAVGSASASYPSPGGGVFGISPP
ESSPRDRKRKRDRPSLVKHTTFPHFDGLWRKYQKNIKDASAFRCSYRDRCLAS
KLVGQENDDPPLRQYNTTPTVPTPDVAOPPPGAGDYLALRFGSSAGG
KGGQAHQOQTPRQQQNTARRPFMLSDSSVHQHLEQHPAFPDGQULPATYAAA
SPSSFTAAELAPPLTTTMDNGDLFSTWDALRYGLDYDHGLGNHVLYLPDCCNGD
DNY"
/complement(join(41449..41706,42198..42595,42641..42848))
/gene="P0703B11.9"
/complement(join(41449..41706,42198..42595,42641..42848))
/notes="contains EST AU094778(E31622)
unknown protein"
/codon_start=1
/protein_id="BAB85290.1"
```

```
/db_xref="GI:18844820"
/translation="MDGDGYGDSGTRDFMSQSDSHADHDELDPYTSLYTYGRSGS
ASRTGMAALDNSGGDWPLDGSYEGGHLWGAGRGVGGSGSHRSASQGGGADLP
PVRASVTLGLRSARPAPCRGATAAANSSPYGGRSPAAPAGRGGRGTQATGA
RGGRGRSTTAATYSEVNLDDSDVDAAELQEI EQNHGKVPKPDRAWNTEDTSL
CEIWDQINGKNCVNGVMQKTGMREIIAAYHAATNLHRGSIQINRIRQLKGLWQFIQ
TLRTNTGFNRI"
/join(43679..43849,43967..44029,45309..45425,45617..45793)
/gene="P0703B11.10"
/join(43679..43849,43967..44029,45309..45425,45617..45793)
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB85291.1"
/db_xref="GI:18844821"
/translation="MKEYGYISHOVWOLLMLLAPPGVOLFGGGMKALTRDGLWRL
TMAGPYOHFGRRLSGSLSWFRGAYLAAGVAYNSRAQITLLDHOTIDROVGEKL
TPSHHLLTFCMFRIQIRHIIKAGLLWTYSLFFEFYAFADIAMLLSSFYGTGSELSL
VSSGSEYKSKSTDTSV"
/join(46353..46642,46739..46876,47886..48585)
/gene="P0703B11.11"
/join(46353..46642,46739..46876,47886..48585)
/notes="hypothetical protein"
/codon_start=1
/protein_id="BAB85292.1"
/db_xref="GI:18844822"
/translation="MAFGQDTIEOLYRELAGRRLSAKLOALEGLPLDSRGOKEAVDV
SRELGRVMVSLYMLKPCSNSSREPGVTRTAPETRTDDSI CLHTPARVKVRSEVL
VRNGREBVRTTEIITSPYKDXGQWRKYQGNKIQDSNYLRLFKCFSRERSCAKX
QVODRAGEPMLVLYNHEHTCQOPQAPGTPNTAGSSPTTTRQROSSSPFAEML
DLTMNAGLSRLLLPHAVGGGSAEAEAAIVTCLAAVISGGGAAAPPLIWPSTA
PEAAFVASAGHSPSADESVADEAAAQWADMDFCFQYQDSTFGAAAAADHRVLIG
DGDGVQIRIVAARIADTVPRYTRDTSWETAGTSSMRGSID"
Query Match 7.5%; Score 107.6; DB 8; Length 144074;
Best Local Similarity 62.0%; Pred. No. 6.6e-10;
Matches 170; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 853 ACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCATGATAAAGC 912
DB 4806 ATAACTTAAGCCCTTCGATCTGAGTTGAGACACAGATTTTCAGCGCACGCAAAACG 4865
QY 913 AGAAAGCTATTAGCACATTAATTACTTAGATATTATTAATAAATTTGAAAAAATAT 972
DB 4866 AGAAAGCTATTAGCATATGATTAATTAAGTATTAAATTTATATAAATTTGAAAAATGGGT 4925
QY 973 TTATTTGAATTTTAAACAATGATGCATAAATTTATTTTAAACACACACCAATTTAA 1032
DB 4926 TTATTTGAATTTCTTAGACAATCTTCTAGTAAACAATTTTGTCAAAGAACCACCAATTTAA 4985
QY 1033 CCTTTAAAAAGCATCCTTAATAGGAAACGAGGAAGTTAAAGATTCCACCGAAGTGTTTGA 1092
DB 4986 TAAATTTAAAAAGCATGTTTAAACGAAACGAGGAGTTGAGAAAAAGCACTGGGSCCTTAGTC 5045
QY 1093 TAAAGAAAAATGGGGTGGGATTAGAAATTCGTTAA 1126
DB 5046 AAACCTTGCTAAGTTGACTTTAAAAAATAGTCAA 5079
RESULT 12
AB023482
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
clone P0680A03.
ACCESSION
AB023482
VERSION
AB023482.2
GI:5091496
KEYWORDS
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0680A03.
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

| | | | |
|---|--|-----|---|
| REFERENCE AUTHORS TITLE | Erhartoideae; Oryzae; Oryza. 1 Sasaki, T., Nagamura, Y. and Yamamoto, K. clone: P0680A03 | CDS | DDGFEQEENLTQLYGLSPGGHCKGNKCNENIPLPQSGDAGLRCHWSLHPQNGRVAPS 1PRNGRAPSPLPCSLPRATRDVTVPRGGMVVRDRRGIFV" complement (join(8393, 8800, 9974, .10171)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78735.1" /db_xref="GI:5091500" |
| JOURNAL REFERENCE AUTHORS TITLE JOURNAL | 2 (bases 1 to 156054) Sasaki, T., Nagamura, Y. and Yamamoto, K. Direct Submission Submitted (05-FEB-1999) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tsasaki@agr.affrc.go.jp, Tel: +81-298-38-7441, Fax: +81-298-38-7468) | CDS | /translation="MQSIIVMSCRHVTEAGLVAVGGCRRLECNVGMVPPPSFAG LLAIRPALQIRSIPOILNANVQOGGAAGAKRNLGLAARDRSRSGAGRWRSQTATVPH TGRAVPRTRITAVPHRGGLDRACATGLDRPCGTAVDCGTGKGLGPARWVCGTV AVPEGAVARAGSAEPWRSRALVLRNRSRAHWCGLCVATV" complement (join(11764, .11781, 12101, .12236, 12390, .12536, 12622, .12787, 13396, .13751, 14014, .14018)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78736.1" /db_xref="GI:5091501" |
| COMMENT | On Jun 17, 1999 this sequence version replaced gi:4521152. Detailed information of this PAC DNA sequence is available at http://www.dna.affrc.go.jp:82/ The orientation of the sequence is from SP6 to T7 of the PAC clone. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN1.4, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against the non-redundant database NRP (PIR, SWISSPROT, GENPEPT, PDB) from MAFF DNABank and the cDNA sequence database at RGP. Protein similarities of the coding regions were searched against NRP with BLAST2.0. ESTs represent the identified cDNA sequences using BLASTN1.4 with the corresponding DDBJ accession no. and RGP clone ID. Sequence updated (08-Mar-1999) Sequence updated (10-Jun-1999). Location/Qualifiers 1. .156054 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="6" /clone="P0680A03" join(1215, .1653, 1694, .1973, 2316, .2405, 2407, .2869) /note="Similar to Arabidopsis thaliana mRNA for DREB1B, complete cds.(AB007788)" /codon_start=1 /protein_id="BAA78732.1" /db_xref="GI:5091497" | CDS | /translation="MESARGRCSRPEWRGGEAGGGGGGARRGAVAAEPHLRSGPS HEEEVGRGGRRGTTEELRFTTLRRPRRSATCGGEMRWSTAEFPVOGRSGAPL LRGCVLHLELTWTRLTARGEATEPKRAGAGVGVVAPPAABEPEPHYERLPADL LAHVLSLLPSPHDLISMAGVGRWRRAVERSLAGRMRMSFAGQRTGDDSTARFVRAAV NLRDLDIRSCWGCOTDQGLIRISTADCVKNLTSISLWGLAGITDNGVIOLCIRSQ" join(14440, .14578, 18996, .20374) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78737.1" /db_xref="GI:5091502" |
| FEATURES source | /translation="MSGGAATVAPGWRRELAGLLVPLVSAMEARGCGHHRRDLDA SAIKLSVMTVRSQGRQRRYARMSPSPPTLNCRLRRTRHGEHRQRAI DDLRLGGDAVVAEAGANOVGDEAPVPMGAMVGVGVOCTGDTAEETARTHD RAMALCAGASALNPAADSAWLLHVPRAPVVSGLMPAARCATCLOGHRVPAFGRS TATATATGDAASTAPPAPVLSSHASSSMLATSVQNLRLATSSHLSPSHERTMP SGAESLAPLLPFRSARLLSPPLHRARPTTPRPAVLFPFHVSAASALLPHRQA VKPPCSHVAAAMQVETGAPGGGAQRSPPIGWSLSLPPAAALLPQPSVHRHSP AGCAPREKSRGKERRRERGWERGSADVAS" join(4728, .4859, 4958, .5037, 5136, .5280, 5374, .5799) /note="ESTs AU058067(E20733), AAU058070(E20873) correspond to a region of the predicted gene Similar to Populus tremuloides caffeoyl-CoA 3-O-methyltransferase mRNA, complete cds.(U27116)" /codon_start=1 /protein_id="BAA78733.1" /db_xref="GI:5091498" | CDS | /translation="MDNHAVALQOQDTPRPRARASVSIQGGGPGFWADHVALPR PRCHRPPTPCRRPLPAPASSIHPRAVLPSGAPPHVSAGADTGAAGLAPPHS FQEKKPKPRHAGVASPRIALAPLNSARMORREATVLPPLPPPTPOQPOOAAA VPAVGGGGGGGROYRVNRKWKGVAEIREPNKSRIMLGLSYSTAVARAAYDTA VFYLKRSARLNFDDLDGAGGAGAGAEHRELTAAVIRKKAEEVGARVDAQHSV VGAAAPVLPQPQPPPPQRRRTKNPDNREPPTDSDDE" complement (join(33393, .33435, 33541, .33634, 36278, .36468, 36579, .36833, 36927, .37241, 38176, .38432)) /note="ESTs AU033170(S4339), D41681(S4339) correspond to a region of the predicted gene. Similar to Pisum sativum mRNA for pIAA4/5.1.(X68215)" /codon_start=1 /protein_id="BAA78739.1" /db_xref="GI:5091504" |
| CDS | /translation="WRPHISSLLPSSLCPLPILLSLSRLGKQRYTGIGRWDDRS GARAPAGRRGADCGWEERSARVRRIMRGNHYPWLLPPTGARAPLLRSPLM SPLGSGTTPPRGQRLRRHSHPAPLPHSPHPPHRRSGSGRGGCGGRLAFLP LUPPTKTTTMTAGSRQRLRRSSQPPPPPSLSLSTPSNVDSIIAMGNATAEDDD | CDS | /translation="MGLWKNRRRDMRDGKGKGSNNDOIPTGNRRDEDOELPSCCTASH GNERSARGORTHTADTAPTRHDCGRGRKREPPIHRASRGSRVRDARLDLSGLKT TFTNPARTLTLHLPLPLNSNR" |

| | | | |
|---|--|----------|--|
| CDS | join(45269. .45340,45973. .46117,46123. .46571) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78741.1" /db_xref="GI:5091506" /translation="MRRFNQNNFEYVLAFILIRPYPTFAEGCKSVNLKSTSTILRD SSRFAASAQPLLPFFGCCPOTKRSALRASKAATSHRSPMSPPHPSPAPRSPF HRLLRPPRPVPIRRPQGVLPAYVSDTAALRSPNPYDNKSGNGASADCAFH LLDEIGLRQRTTPASSRAPRSLTLARAHGSPCWTLRPRMFRPTWRARCSGCTSP VIV" complement(join(48086. .48195,48370. .48525,48574. .48616)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78742.1" /db_xref="GI:5091507" /translation="MODLPQVLNVVFGEAAARLVBESGATGCGCGGTEKAVGGG AARCSLAVGRRRGAATQORRKEGTVNDYPAVSTYPPYDAYQILLIGSAVSG " join(48938. .48993,51122. .51322,52056. .52200,52695. .52957, 53080. .53704) /note="ESTs C22458(C62866),C22459(C62866) correspond to a region of the predicted gene. Similar to genomic sequence of Arabidopsis thaliana BAC F8A5, complete sequence. (AC002292)" /codon_start=1 /protein_id="BAA78743.1" /db_xref="GI:5091508" /translation="WVTPFGIAVSRIRIRINVSCKSSBSTSLSCSTTTTSSFRQLFFP FVSGAARLDLGRNAAEMPKLSRFAHRVLCGRKSGEDLSDEGSGSLRWVFSLE LRATNSFNDKIGEGFSGVYMGQSGQIAVKLCKAKNGTETEFASDVEILGR VRHKNLUSFRYGCADGPBRLVDFMPSNLQHLGHSTECLLDWRRTFFIAIGAA RALAYLHHHTPQIHGSKATNVLLDSNFQHLGDFGLIRFIPDGVHDKIIENOR GYLAPEYFMFKPTIGCDVSGFIIILLESRRPVERSGSAKCVGNWVLPKADG RYDELVDSKLDNKYSESLKRVLVGLACTHREPEKPTMLEVVSMLKGESKMSRL ENDELFRPDS"TVSSHGNSTPGSSDCVCPKNDQSLAAA" complement(join(54288. .54381,54733. .54920,55012. .55254, 55374. .55652,55743. .55940,56029. .56664,56763. .57017, 57133. .57239,57324. .61120,61231. .61720,62472. .62862, 63035. .63127,63241. .63364,64098. .64378)) /note="EST AU065533 (C21174) corresponds to a region of the predicted gene. Similar to Homo sapiens splicing factor Prp8 mRNA, complete cds. (AF092565)" | ORGANISM | Oryza sativa Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoae; Oryza. 1 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Henry, D. and Simmons, J. Rice Genomic Sequence Unpublished 2 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Henry, D., Simmons, J., Wilson, R., Minx, P., Bradshaw, H. and Du, H. Direct Submission Submitted (17-MAY-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA 4 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H., Rambo, T., Henry, D., Simmons, J., Wilson, R., Minx, P., Bradshaw, H., Du, H., Thurmond, S.K. and Mao, L. Direct Submission Submitted (24-MAY-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA On May 17, 2001 this sequence version replaced gi:13384341. This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=30: an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone: and the assembly was confirmed by restriction digest. Location/Qualifiers 1. .129838 /organism="Oryza sativa" /db_xref="taxon:4530" /chromosomes="7" /clone="OSJNBb0024A20" 1713. .3591 /gene="OSJNBb0024A20.1" /note="Putative Mutator protein" /protein_id="AAK55767.1" /db_xref="GI:14192862" /translation="MDRLVRLHYGGSVVQSNNGNFGMSVKQLIFGKPSFEEELVC HTKDLVQNSQSIATQGRYDVAGPI SHKYMLDLNGLEWNTYVDIVLGSQFSLLEV ASKLDRDRCDDSFDLNRTLFYDEPNYETFDDEGDMKSQSVKSLTVEEVIESQMGVD VNMKCEADAKGRSQRSEVEINVELQKEEAANKASKLES939SVNVELQKEEA IEVNEARQAPVINGNVDNDDNEVDAGSMGGREERTTDAQLASLVRFCFRDKIYTRR ECEETLLMGLSVDLLHESEYDDTHEEGYEYVAVSDDRPVAKISQREBEVFE KLVCNPEVQFEDVNNSELAVDGVYNGIIDAMGEDIIWHIIFKLSLRLCLLLK RKSKKITDWAHHLHLIERDNTAANVPHGVHKNKTYFDEYELAWLLRNRYLFLRPA WTLADIATDPEVEEMNEYDTHTRAGATVEYGPVDRVARELMSRVNDAGVALGTGPA SENEVGILRSALQRLKQRSRLAARLCGRSTNAVDPPQLRTTQAGETSQARAEESG" 1713. .3591 /note="Putative Mutator protein" complement(10720. .11542) /gene="OSJNBb0024A20.2" /note="Hypothetical protein" complement(join(10720. .11360,11455. .11542)) /gene="OSJNBb0024A20.2" /codon_start=1 /product="Hypothetical protein" /protein_id="AAK55768.1" /db_xref="GI:14192863" |
| FEATURES | source | | |
| gene | | | |
| CDS | | | |
| Query Match 7.4%; Score 106.4; DB 8; Length 156054; Best Local Similarity 77.2%; Pred. No. 1.1e-09; Matches 142; Conservative 0; Mismatches 41; Indels 1; Gaps 1; | | | |
| Qy 893 TTTCAACGCATGATAACGACGAAGCTCATTTAGCACATTATTACTAGATATTATTAAT 952 | | | |
| Db 105226 TTTTTAGCGCAGCGGAAACGAGAAAGCTCATTTAACACATGATTAAATAGTATTGATAAT 105285 | | | |
| Qy 953 TATAAACTTGAATAAATAATTTATTGAAATTTTTTAAACAAATGTATGCATAAATATTATTT 1012 | | | |
| Db 105286 TATAAACTTAAATAATGGAATTTATTGAAATTTTTTAAACAACTTCTATATAGATT-TTTT 105344 | | | |
| Qy 1013 TTAATAACACACCAATTTAAACCTTTTAAACCATCTCTTAATAGGAAACGAGGAAGTTAAA 1072 | | | |
| Db 105345 TTAATAAAGATATCATTTAAACAGTTTGAACCGCTGCTAACGCAAAATGAGGAAGTTGAA 105404 | | | |
| Qy 1073 GATT 1076 | | | |
| Db 105405 ATTT 105408 | | | |
| RESULT 13 | | | |
| AC079038/c | | | |
| LOCUS AC079038 129838 bp DNA linear PLN 24-MAY-2001 | | | |
| DEFINITION Oryza sativa chromosome 7 clone OSJNBb0024A20, complete sequence. | | | |
| ACCESSION AC079038 | | | |
| VERSION AC079038.3 GI:14140271 | | | |
| KEYWORDS HTG. | | | |
| SOURCE Oryza sativa. | | | |

| | | | | | | | |
|---------------------------|--|--|--|--|--|--|--|
| CDS | /gene="OSJNBa0052021.7" | | | | | | |
| | complement(join(26051..27522,28362..28860)) | | | | | | |
| | /gene="OSJNBa0052021.7" | | | | | | |
| | /codon_start=1 | | | | | | |
| | /db_xref="GI:21740458" | | | | | | |
| | /translation="MCPGQSDSNKQKRWTLNRQGLGLOQEIKHASLFCGVSHGL WRAALPQRIFFRVTDCAVIGDICEDEPHGLAPKIPSSNSRSLVHLPPYRGKISS RYANRPGKGRGKSLGKNGVLSRRHFATAASPAIQTFRRLDTSSTSPSP SSPAGEGGGFMIPTAKVSGRPFESKATVRVNGHISIDNLTTELRLVPRFE VKETAQDTFVVQFSTIELDLHRVGRVSFDFGVLLFNKWEQIEAKVLKKTWR VYGVPLRDLJHVASVIGNTTMDMTVLRKHGVVRMCVADDDNLTLPSPYIVF RNHYEIFFEEMPEESKGEKLEDDQSDSKRNDVDDVMDNRLKLNAAAS SSMLRGSGVTDGSLNHCVCISAIAGLLTVDDSSAVEQOAEVLPTICVASSLETVMNK APESSSMOIPTTGQNTLANEAVALLVPDVAALISBEFTTLDGEMAPAGCA SLSPASTGATLEGAVTVSLAAPVVMDDAAQIEESDKVGLVAADAPLTSVDATCVLPS SSTGTMDAAPNWEVEVTEVTSQSMILDDVGTAGVNTVIELAADGAQDAARTIIIEA KVQAEAGGKSPQAKRSRLIHPITRQSE" | | | | | | |
| gene | complement(29690..30004) | | | | | | |
| CDS | /gene="OSJNBa0052021.8" | | | | | | |
| | complement(29690..30004) | | | | | | |
| | /gene="OSJNBa0052021.8" | | | | | | |
| | /codon_start=1 | | | | | | |
| | /protein_id="CAD40023.1" | | | | | | |
| | /db_xref="GI:21740459" | | | | | | |
| | /translation="MEVDAACVEEALGAVFAVGAEIRITQPVLVHATGADADGHVGR GDPAAGRAAVGGDSAPVSSGAGTGGAAHGVAATAAPADATENEVWATKAANETA AK" | | | | | | |
| gene | 39190..45154 | | | | | | |
| CDS | /gene="OSJNBa0052021.9" | | | | | | |
| | join(39190..40045,40272..41545,41618..45154) | | | | | | |
| | /gene="OSJNBa0052021.9" | | | | | | |
| | /codon_start=1 | | | | | | |
| | /protein_id="CAD40024.1" | | | | | | |
| | /db_xref="GI:21740460" | | | | | | |
| | /translation="MAERAVAHNLSPSAGSDGQEQNRPRRARTPPSPNLRGGEAL EGVRSATSPGAGGGRDNRRLVYGDGTQGALQAGALLRPVPPVPEPTA RRLDLDVAVLQOALGAGSSANKASGAAAGSTSSRRARRAAAVTRSSAIP SSTPTQDVHGGDARLDIRERRDDRAAHATGASGSSGAPLRSHGGRPPVSDV AGCAFVASLNRWMPFRPTITEKYDGSVNPAEFLQIYTTGIEAAGGDRVMANFF PMALKGAGGGVRHNRLEKIASKEPOTTAQLFELADRVARKEEATWSSSGVAA PAAFGSAGGRRRRKRSKSDDEGHVLAEGASRPKRPSPADKKKAGAPRGE ERPTKWTCHTSLDLADYAVKSLAERTRWEEKQREKGAAPAPAGNRGE AKQRTADIDGNDLGFQEPFATVATDGGCAHAHSRRSLKMKELLAALITLSEA TRARWSVALTFQIDHPPCVARGGOIAMVSPVCNKLGRVLIQDGAALITLSPV AFDAIKASGVMLRPSQRIIGVTPGHTWPLGHIDLPVTEGGSANFTEERVDVADLSL PYNVLGRPALVMAVHYAYLQKMPGPGGPIVHGDLKVALACMEQADHLAVAS KPADGERLSTSVATPRQORVTCDEVPEDALVFLRANADVFAWRPADMPGVPREVI EHLAVRFGARPVQRVQRPQAFICEVARLLEAGFTIREVHPHELANPVVVPK ANGKLRMCIDYDLNLCACPKDPYPLRIDQIVDSTAGDGLLCFLDAYSQVHQIWMARE DEEKTAFTITPTGYTTPMPLGNAGPTFORTTRISLGSOLGRNVEAYVDDLVKTR NQETLLSLAEFTFLSARILKLPNKCFCVGPAGKLLGLFVLSARGIEANPEKIRALE RMRPPSKURDQVCTGCMVAULSRFISRLGKALFLKLLKSGPFTWTEEAERAFTEL | | | | | | |
| Query Match | 7.2%; Score 103.6; DB 2; Length 170025; | | | | | | |
| Best Local Similarity | 75.8%; Pred. No. 3.3e-09; | | | | | | |
| Matches 141; Conservative | 0; Mismatches 44; Indels 1; Gaps 1; | | | | | | |
| 893 | TTTCAACGCGATATAACGAGAAAGCTCATTTAGACATTATTACTTAGATATTTATTAAT 952 | | | | | | |
| 107584 | TTTTTGACGCAAGCAACGAGAAATCTCATTTAGCATAGATAATTAAGTATAATAAT 107525 | | | | | | |
| 953 | TATAAACTTGAAAAATAATTTATTGATTTTAAACAAATGATGATAATTTT 1012 | | | | | | |
| 107524 | TATAAACTTGAAAAATAGGTTTATTGAAATTTTAAATAAATCTTATATAATTTT 107465 | | | | | | |
| 1013 | TATAAAACACACCAATTTAAACCTTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAA 1072 | | | | | | |
| 107464 | TGCAAAATGATC-ATTTAACAGTTTGAAGAAGCTGCTTACCGAAACGAGGAAGTTGAA 107406 | | | | | | |
| 1073 | GATTCA 1078 | | | | | | |
| 107405 | GTTTGA 107400 | | | | | | |

| | | | | | |
|------------|---|-------------|-----|--------|-----------------|
| RESULT 15 | AC023240 | 131983 bp | DNA | linear | PLN 29-AUG-2001 |
| LOCUS | AC023240 | | | | |
| DEFINITION | Oryza sativa chromosome 10 BAC OSJNBa0051D19 genomic sequence, complete sequence. | | | | |
| ACCESSION | AC023240 | | | | |
| VERSION | AC023240.8 | GI:15341601 | | | |
| KEYWORDS | HTG. | | | | |
| SOURCE | Oryza sativa | | | | |
| ORGANISM | Oryza sativa | | | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza. | | | | |
| REFERENCE | 1 (bases 1 to 131983) Buell,C.R., Yuan,Q., Moffat,K.S., Hill,J.N., Jenkins,C.N., Burr,P.C., Hsiao,J., Zismann,V., Pai,G., Bowman,C.L., Fujii,C.Y., Vankken,S.E., Bowman,C.L., Craven,B., Utterback,T.R., Khalak,H., Feldblyum,T.V., Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M. | | | | |
| AUTHORS | Oryza sativa chromosome 10 BAC OSJNBa0051D19 genomic sequence | | | | |
| TITLE | Unpublished | | | | |
| JOURNAL | 2 (bases 1 to 131983) | | | | |
| AUTHORS | Buell,R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (10-FEB-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA | | | | |
| REFERENCE | 3 (bases 1 to 131983) | | | | |
| AUTHORS | Buell,R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (15-SEP-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org | | | | |
| REFERENCE | 4 (bases 1 to 131983) | | | | |
| AUTHORS | Buell,R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (14-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org | | | | |
| REFERENCE | 5 (bases 1 to 131983) | | | | |
| AUTHORS | Buell,R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (29-AUG-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuelli@tigr.org | | | | |
| COMMENT | On Aug 29, 2001 this sequence version replaced gi:10140693. Address all correspondence to:rice@tigr.org | | | | |
| | BAC clone OSJNBa0051D19 is from Oryza sativa chromosome 10 | | | | |
| | The orientation of the sequence is from SP6 to T7 end of the BAC clone. | | | | |
| | Genes were identified by a combination of several methods: Gene prediction programs including Genscan and Genscan+ (Chris Burge, http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky, http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by repeatmasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html). | | | | |
| FEATURES | Location/Qualifiers | | | | |

RESULT 15

AC023240

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

AC023240 131983 bp DNA linear PLN 29-AUG-2001
Oryza sativa chromosome 10 BAC OSJNBa0051D19 genomic sequence,
complete sequence.

AC023240
AC023240.8 GI:15341601
HTG.

Oryza sativa.

ORGANISM

Oryza sativa.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

1 (bases 1 to 131983)

Buell, C.R., Yuan, Q., Moffat, K.S., Hill, J.N., Jenkins, C.N.,
Burr, P.C., Hsiao, J., Zismann, V., Pai, G., Bowman, C.L., Fujii, C.Y.,
Vanaken, S.E., Bowman, C.L., Craven, B., Utterback, T.R., Khalak, H.,
Feldblyum, T.V., Quackenbush, J., White, O., Salzberg, S.L. and
Fraser, C.M.

Oryza sativa chromosome 10 BAC OSJNBa0051D19 genomic sequence

Unpublished

2 (bases 1 to 131983)

Buell, R.

Direct Submission

Submitted (10-FEB-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

3 (bases 1 to 131983)

Buell, R.

Direct Submission

Submitted (22-SEP-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

5 (bases 1 to 131983)

Buell, R.

Direct Submission

Submitted (14-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

6 (bases 1 to 131983)

Buell, R.

Direct Submission

Submitted (29-AUG-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

On Aug 29, 2001 this sequence version replaced gi:10140693.
Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0051D19 is from Oryza sativa chromosome 10

The orientation of the sequence is from SP6 to T7 end of the BAC

clone.

Genes were identified by a combination of several methods: Gene

prediction programs including Genscan and Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer

(Mihaila Parva and Steven Salzberg, contact mparva@tigr.org),

searches of the complete sequence against a peptide database and

the plant EST database at TIGR (http://www.tigr.org/tdb/cgi.shtml).

Annotated genes are named to indicate the level of evidence for

their annotation. Genes with similarity to other proteins are named

after the database hits. Genes without significant peptide

similarity but with EST similarity are named as unknown proteins.

Genes without protein or EST similarity, that are predicted by more

than two gene prediction programs over most of their length are

annotated as hypothetical proteins. Genes encoding tRNAs are

predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are

identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RN/RepeatMasker.html).

Location/Qualifiers

Db 78507 ATTCATTGGAA-TTTTAAACAACCTTTTACATAAATTTTTTATATAAAACACACTATTT 78565

Qy 1031 AACCCTTTAAAAAGCATCCTAATAGGAAACGAGGAAAGTTAAAGATT 1076

Db 78566 AACCGTTTGAAAAAACGATGATAATGGAAACCGAGGAAAGTTGTAATTT 78611

Search completed: March 15, 2003, 15:53:37
Job time : 5046.28 secs

PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
PT induction in recombinant plants -

XX Claim 10; Page 28; 33pp; Japanese.

XX This sequence represents a fragment of the rice adenylate kinase gene.
CC The invention relates to a DNA fragment (such as this sequence), that has
CC a regulatory function (preferably acting as a promoter) on the expression
CC of desired foreign structural genes in a plant when introduced into a
CC plant body or plant cell. The DNA fragment can be inserted into a plant
CC to control the expression of desired foreign structural genes, useful for
CC inducing in the plant e.g. herbicide, stress and light resistance,
CC production of antibacterial and insecticidal substances, sucrose
CC production of recombinant plants, and in promoting flowering and crop
CC yield. The DNA fragment has a regulatory function on the expression of
CC desired structural genes in a plant, particularly with control of the
CC expression in specific tissues and locations to provide safe chemicals.

XX SQ Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

Query Match 100.0%; Score 1440; DB 21; Length 1440;
Best Local Similarity 100.0%; Pred. No. 9.6e-288; Indels 0; Gaps 0;
Matches 1440; Conservative 0; Mismatches 0;

QY 1 CTGCAGGAAGATTAAATTAGGTGGACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60
DB 1 CTGCAGGAAGATTAAATTAGGTGGACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60
QY 61 AATCAATGGGGTGGTGGTGGACATGGTTTTGCGAGGAAATTAAGCAAGAAATTAAG 120
DB 61 AATCAATGGGGTGGTGGTGGACATGGTTTTGCGAGGAAATTAAGCAAGAAATTAAG 120
QY 121 AAGAAATGCTCAAGCTGACATGAGAAACGTAATCCAAATGGAACGGAATTTCAAGTCGTTTC 180
DB 121 AAGAAATGCTCAAGCTGACATGAGAAACGTAATCCAAATGGAACGGAATTTCAAGTCGTTTC 180
QY 181 TCTTGACTACCTGACTGTGTCACAAAGCATCAATGCTTCTCGAGTATCTTTATTA 300
DB 181 TCTTGACTACCTGACTGTGTCACAAAGCATCAATGCTTCTCGAGTATCTTTATTA 300
QY 241 CTGCTGTGATACCTGACTGTGTCACAAAGCATCAATGCTTCTCGAGTATCTTTATTA 300
DB 241 CTGCTGTGATACCTGACTGTGTCACAAAGCATCAATGCTTCTCGAGTATCTTTATTA 300
QY 301 CCGAAATCCCAAGATTAATTTCTTCCACCTCAGGGTAATTTGCTGAACTATGCAATG 360
DB 301 CCGAAATCCCAAGATTAATTTCTTCCACCTCAGGGTAATTTGCTGAACTATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGGTATCTATCTTCTGCTCAAAATTTGAGTCCAA 420
DB 361 AATACAAATTCGCAAAATATCATGGTATCTATCTTCTGCTCAAAATTTGAGTCCAA 420
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480
DB 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480
QY 481 CGCAATTCACCGTTTCGAGAAATATGCTGTATATAAATAGTAGTCTAGTGCAGAAACAA 540
DB 481 CGCAATTCACCGTTTCGAGAAATATGCTGTATATAAATAGTAGTCTAGTGCAGAAACAA 540
QY 541 ATTAATATACATATAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTACACTCT 600
DB 541 ATTAATATACATATAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTACACTCT 600
QY 601 AATTTGTAAATTTCTTATTTTCAGTACAAAATTCCTCAATTTTCAATTAAGAAATAAAGCT 660
DB 601 AATTTGTAAATTTCTTATTTTCAGTACAAAATTCCTCAATTTTCAATTAAGAAATAAAGCT 660
QY 661 AGACGGCTAAGCCCAACCTCTAAGGCTAAGTTCGAGAGGTGAAGTACGACGCAAAATA 720
DB 661 AGACGGCTAAGCCCAACCTCTAAGGCTAAGTTCGAGAGGTGAAGTACGACGCAAAATA 720

QY 721 TGATGGTTTAAATATATGATTTTAAATAAATTTTACATAAATTTCTTTAGGAAACA 780
DB 721 TGATGGTTTAAATATATGATTTTAAATAAATTTTACATAAATTTCTTTAGGAAACA 780
QY 781 TATCATTTAAATGGTTTGAATAACCGTGCAATAGAAACATAGATTTGGGAAA 840
DB 781 TATCATTTAAATGGTTTGAATAACCGTGCAATAGAAACATAGATTTGGGAAA 840
QY 841 CAAGAGAAACACACAGCTTAAAGCTTCTTCTAGTTGGAGGTTGATTTTCAAC 900
DB 841 CAAGAGAAACACACAGCTTAAAGCTTCTTCTAGTTGGAGGTTGATTTTCAAC 900
QY 901 GCATGATAAAGCAGAAAGCTCATTTAGCACATTTACTTAGATTTTATAATTAATAACT 960
DB 901 GCATGATAAAGCAGAAAGCTCATTTAGCACATTTACTTAGATTTTATAATTAATAACT 960
QY 961 TGAAAAAATATTTTAAATTTTAAACAAATGTATGCATAAATTTTAAAAAC 1020
DB 961 TGAAAAAATATTTTAAATTTTAAACAAATGTATGCATAAATTTTAAAAAC 1020
QY 1021 ACACCAATTTAAACCTTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTCAACC 1080
DB 1021 ACACCAATTTAAACCTTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTCAACC 1080
QY 1081 GAAGTGTGGATAAATGAATAATGGGTGGGATAGAAATGGTAAATGAATCAGGGTTAG 1140
DB 1081 GAAGTGTGGATAAATGAATAATGGGTGGGATAGAAATGGTAAATGAATCAGGGTTAG 1140
QY 1141 GATTAAATATTTAAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTTTTGG 1200
DB 1141 GATTAAATATTTAAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTTTTGG 1200
QY 1201 GTGGGTAGAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCGTGTGGGTGGGCGAGA 1260
DB 1201 GTGGGTAGAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCGTGTGGGTGGGCGAGA 1260
QY 1261 AGCGTCTTTTGGTGGAAAAAATCTGCTTAAAGAGAAACAGAAAGAGCCGAGCTTT 1320
DB 1261 AGCGTCTTTTGGTGGAAAAAATCTGCTTAAAGAGAAACAGAAAGAGCCGAGCTTT 1320
QY 1321 GTTGTACCGTCTCACAGAAACGAAACAAAAAGCCCCCACCACCTTAAACCTCTCGATCC 1380
DB 1321 GTTGTACCGTCTCACAGAAACGAAACAAAAAGCCCCCACCACCTTAAACCTCTCGATCC 1380
QY 1381 GACCGAGACTCTCTCATTTTACGCGCGCACGCGGAGAGACGCGAGTTCGTCGAC 1440
DB 1381 GACCGAGACTCTCTCATTTTACGCGCGCACGCGGAGAGACGCGAGTTCGTCGAC 1440

RESULT 2

AAAA10961
ID AAA10961 standard; DNA; 1440 BP.

XX AAA10961;

XX AC AAA10961;

XX DT 14-JUL-2000 (first entry)

XX DE Partial rice adenylate kinase gene promoter fragment.

XX KW Rice; adenylate kinase; promoter; gene expression control; antibacterial;

XX KW insecticide; light resistance; sucrose induction; flower promotion;

XX OS Oryza sativa.

XX PN WO200015811-A1.

XX PD 23-MAR-2000.

XX PF 10-SEP-1998; 98WO-JP04088.

XX PR 10-SEP-1998; 98WO-JP04088.

XX XX

PD 23-MAR-2000.
XX 08-SEP-1999; 99WO-JP04847.
XX 10-SEP-1998; 98WO-JP04088.
XX (NISC) NISSAN CHEM IND LTD.
XX Uchimiya H, Arai S, Fushimi T, Tagawa M, Fukuzawa H;
PI WPI; 2000-271448/23.
DR
XX
XX Rice adenylylate kinase gene-originated DNA fragment with promoter
PT function in plants for controlling expression of desired foreign
PT structural genes, useful for herbicide resistance and in sucrose
PT induction in recombinant plants -
XX
XX Claim 1; Page 25-27; 33pp; Japanese.
XX
XX This sequence represents a fragment of the rice adenylylate kinase gene.
CC The invention relates to a DNA fragment (such as this sequence), that has
CC a regulatory function (preferably acting as a promoter) on the expression
CC of desired foreign structural genes in a plant when introduced into a
CC plant body or plant cell. The DNA fragment can be inserted into a plant
CC to control the expression of desired foreign structural genes, useful for
CC inducing in the plant e.g. herbicide, stress and light resistance,
CC production of antibacterial and insecticidal substances, sucrose
CC induction in recombinant plants, and in promoting flowering and crop
CC yield. The DNA fragment has a regulatory function on the expression of
CC desired structural genes in a plant, particularly with control of the
CC expression in specific tissues and locations to provide safe chemicals.
XX
SQ Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;

Query Match 100.0%; Score 1440; DB 21; Length 4341;
Best Local Similarity 100.0%; Pred. No. 1.2e-287;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGCAGCCCTGTGTT 60
Db 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGCAGCCCTGTGTT 60

Qy 61 AATCACTGGGGTGTCTTGGACATGTTTTTCAGGAAATTAAGCAAGAAATTAAG 120
Db 61 AATCACTGGGGTGTCTTGGACATGTTTTTCAGGAAATTAAGCAAGAAATTAAG 120

Qy 121 AAGAAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTT 180
Db 121 AAGAAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTT 180

Qy 181 TCTTGTAATCAATGTTTGAATATAGACAGTGCACAGTGGTTCCTATG 240
Db 181 TCTTGTAATCAATGTTTGAATATAGACAGTGCACAGTGGTTCCTATG 240

Qy 241 CTGCTGTGATCTGACTGTCTCAAGAGCATCAATTTGCTTCTGGAGTATCTTTATTA 300
Db 241 CTGCTGTGATCTGACTGTCTCAAGAGCATCAATTTGCTTCTGGAGTATCTTTATTA 300

Qy 301 CCGAAACCCCAAGAAATTAATTTCCACCTCAGGGTAATTTGCTGAACTATGCAATG 360
Db 301 CCGAAACCCCAAGAAATTAATTTCCACCTCAGGGTAATTTGCTGAACTATGCAATG 360

Qy 361 AATCAAAATTCGCAAAATATCATGGTATATCTTCTGCTCAAAATTTGAATTTGAGTCCAA 420
Db 361 AATCAAAATTCGCAAAATATCATGGTATATCTTCTGCTCAAAATTTGAATTTGAGTCCAA 420

Qy 421 CTGAGACTGCAATACGATTTTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTCATGAAA 480
Db 421 CTGAGACTGCAATACGATTTTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTCATGAAA 480

Qy 481 CGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCAGCAAAACAA 540
Db 481 CGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCAGCAAAACAA 540

Qy 541 ATTAATATCACATAAAAAAGAGGTGTTTAATTACAAACCATGTTTCGTACTACAACTCT 600
Db 541 ATTAATATCACATAAAAAAGAGGTGTTTAATTACAAACCATGTTTCGTACTACAACTCT 600

Qy 601 AATTGTAAATTTCTTATTTTCAGTCAAAAAATTCGAATTTCCAAATTAAGAAAAATAAAGCT 560
Db 601 AATTGTAAATTTCTTATTTTCAGTCAAAAAATTCGAATTTCCAAATTAAGAAAAATAAAGCT 560

Qy 661 AGACGGCTAAGCCCAACCCATCTAAGCTTAAGTTCGAGAGGTGAAGTACGACGAAAAATA 720
Db 661 AGACGGCTAAGCCCAACCCATCTAAGCTTAAGTTCGAGAGGTGAAGTACGACGAAAAATA 720

Qy 721 TGATGTTTATTAATATGATTTTTTTTAAATTAACCTTTCACATAAAATTTCTTTAGGAACA 780
Db 721 TGATGTTTATTAATATGATTTTTTTTAAATTAACCTTTCACATAAAATTTCTTTAGGAACA 780

Qy 781 TATCATTTAATGTTTGAAAAAACGTGCACATAAGAAAACTAAGAACGATGAGTTCGGAAA 840
Db 781 TATCATTTAATGTTTGAAAAAACGTGCACATAAGAAAACTAAGAACGATGAGTTCGGAAA 840

Qy 841 CAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAC 900
Db 841 CAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAC 900

Qy 901 GCATGATAAACGAGAAAGCTCAATTAGCACATTTACTTAGATATTTAATATTAATACT 960
Db 901 GCATGATAAACGAGAAAGCTCAATTAGCACATTTACTTAGATATTTAATATTAATACT 960

Qy 961 TGAAGAAAAATATTTATTTGAAATTTTTTAAACAATGTATGCAATAAATTTTAAAAAAC 1020
Db 961 TGAAGAAAAATATTTATTTGAAATTTTTTAAACAATGTATGCAATAAATTTTAAAAAAC 1020

Qy 1021 ACACCAATTTAACCCCTTTAAAGAGCATCTATAGGAACGAGGAGTTAAAGATTCACC 1080
Db 1021 ACACCAATTTAACCCCTTTAAAGAGCATCTATAGGAACGAGGAGTTAAAGATTCACC 1080

Qy 1081 GAAGTGTGTTGGATAATGAAAAATGGGGTGGGATTAGAATTTGTAATGAATCAGGGTTAG 1140
Db 1081 GAAGTGTGTTGGATAATGAAAAATGGGGTGGGATTAGAATTTGTAATGAATCAGGGTTAG 1140

Qy 1141 GATTAAATTAATTAAGAGGAGGAGTAATGATGTTAGAGTTTAAATGTGCTTTTGTG 1200
Db 1141 GATTAAATTAATTAAGAGGAGGAGTAATGATGTTAGAGTTTAAATGTGCTTTTGTG 1200

Qy 1201 GTGGGTAGAAAAATTAATTTGCCATACACTCCCGAGAGCGGTGCTTGGCTGGGCAGA 1260
Db 1201 GTGGGTAGAAAAATTAATTTGCCATACACTCCCGAGAGCGGTGCTTGGCTGGGCAGA 1260

Qy 1261 AGGCTCTTTTTCGTTGGAAAAAAAATCTGCTTAAAAAGGAAACAGAAAGAGCCAGCTTT 1320
Db 1261 AGGCTCTTTTTCGTTGGAAAAAAAATCTGCTTAAAAAGGAAACAGAAAGAGCCAGCTTT 1320

Qy 1321 GTTGTCAACCGTCTCACCAGAAAAAGAAAAAGCCCAACCACTTAAACCTCTCGATCC 1380
Db 1321 GTTGTCAACCGTCTCACCAGAAAAAGAAAAAGCCCAACCACTTAAACCTCTCGATCC 1380

Qy 1381 GACCGAGACTCTCCATTTTCAGCGCGGACGCGAGAGACGCGAGCTTCGTCGAC 1440
Db 1381 GACCGAGACTCTCCATTTTCAGCGCGGACGCGAGAGACGCGAGCTTCGTCGAC 1440

RESULT 4
AA10960
ID AA10960 standard; DNA; 4341 BP.
XX
AC AAA10960;
XX
DT 14-JUL-2000 (first entry)
XX
DE Rice adenylylate kinase gene promoter fragment.
XX
KW Rice; adenylylate kinase; promoter; gene expression control; antibacterial;


```

DT 10-JUN-2002 (first entry)
DE Rice Wanderer-Osl nucleotide sequence SEQ ID NO:28.
XX
XX Rice; plant; detection; polymorphism; transposable element; genome;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO200212484-A1.
XX
PD 14-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-JP06661.
XX
PR 02-AUG-2000; 2000JP-0234577.
XX
XX (NISR) JAPAN TOBACCO INC.
PA (SYGN) SYNGENTA LTD.
XX
PI Komori T, Nitta N;
XX
DR WPI; 2002-241759/29.
XX
XX Construction of marker for detecting plant genome polymorphism with use
PT of transposable element, useful particularly in studying restriction
PT fragment length polymorphism applicable in cell genetics -
XX
PS Example 2; Fig 7; 110pp; Japanese.
XX
XX The present invention describes a method for constructing a marker for
CC detecting polymorphisms in a plant genome. The method comprises the
CC production of a primer for nucleic acid amplification by using the base
CC sequences of a transposable element and/or the domain adjacent to it.
CC Also described is a marker for detection polymorphisms in plant genomes.
CC The constructed marker can be used for detecting plant genome
CC polymorphisms, which is useful particularly in studying restriction
CC fragment length polymorphism applicable in cell genetics e.g. for
CC analysing and selecting specific breeds of plants. The method is simple,
CC easy, less time consuming and not so laborious, e.g. in the study of
CC less frequently occurring polymorphism between various species to enable
CC identification of the site and isolation of the required gene.
CC ABL49926 to ABL50036 represent nucleotide sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 422 BP; 145 A; 65 C; 56 G; 156 T; 0 other;

Query Match 6.9%; Score 100; DB 24; Length 422;
Best Local Similarity 70.5%; Pred No. 1,9e-11;
Matches 148; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

QY 869 TTGATCCTCTAGTTGGAGTTGATTTTCAAACGCGATGATAAACGAGAAAGCTCATTAGCA 928
DB 82 TTAATCTCCACATTTGGAGATAAATTTTGGTTGCACGTAACGAGAAAGCTCATTAGCA 141
QY 929 CATTATTACTAGATATTTATATTAATTAACCTTGAAAAAATATTTATTGAAATTTTAA 988
DB 142 CATATTAATTAAGTATTAACCTATTAATAATTTGAAAAATGGAATTTATTTCTTTTAA 201
QY 989 AACAA--TGATGCTAAATATTTTAAACACACAAATTAACCTTTTAAAGCA 1046
DB 202 AACAACTTCTATATAGAACTTTTAAACAAAAATGTATCATTTTAAACAAATTTGAAAGCG 261
QY 1047 TCCTAATAGGAAACGAGGAAGTTAAAGATT 1076
DB 262 TGCTAACGAAAACAAGGAAGTTGAGAAAT 291

RESULT 6
ABL49955
ID ABL49955 standard; DNA; 423 BP.
XX
AC ABL49955;

```

```

XX 10-JUN-2002 (first entry)
XX Rice Wanderer-Osl nucleotide sequence SEQ ID NO:30.
DE
XX Rice; plant; detection; polymorphism; transposable element; genome;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO200212484-A1.
XX
PD 14-FEB-2002.
XX
PF 02-AUG-2001; 2001WO-JP06661.
XX
PR 02-AUG-2000; 2000JP-0234577.
XX
XX (NISR) JAPAN TOBACCO INC.
PA (SYGN) SYNGENTA LTD.
XX
PI Komori T, Nitta N;
XX
DR WPI; 2002-241759/29.
XX
XX Construction of marker for detecting plant genome polymorphism with use
PT of transposable element, useful particularly in studying restriction
PT fragment length polymorphism applicable in cell genetics -
XX
PS Example 2; Fig 7; 110pp; Japanese.
XX
XX The present invention describes a method for constructing a marker for
CC detecting polymorphisms in a plant genome. The method comprises the
CC production of a primer for nucleic acid amplification by using the base
CC sequences of a transposable element and/or the domain adjacent to it.
CC Also described is a marker for detection polymorphisms in plant genomes.
CC The constructed marker can be used for detecting plant genome
CC polymorphisms, which is useful particularly in studying restriction
CC fragment length polymorphism applicable in cell genetics e.g. for
CC analysing and selecting specific breeds of plants. The method is simple,
CC easy, less time consuming and not so laborious, e.g. in the study of
CC less frequently occurring polymorphism between various species to enable
CC identification of the site and isolation of the required gene.
CC ABL49926 to ABL50036 represent nucleotide sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 423 BP; 146 A; 66 C; 56 G; 155 T; 0 other;

Query Match 6.8%; Score 97.4; DB 24; Length 423;
Best Local Similarity 69.7%; Pred No. 6.4e-11;
Matches 147; Conservative 0; Mismatches 61; Indels 3; Gaps 1;

QY 869 TTGATCCTCTAGTTGGAGTTGATTTTCAAACGCGATGATAAACGAGAAAGCTCATTAGCA 928
DB 82 TTAATCTCCACATTTGGAGATAAATTTTGGTTGCACGTAACGAGAAAGCTCATTAGCA 141
QY 929 CATTATTACTAGATATTTATATTAATTAACCTTGAAAAAATATTTATTGAAATTTTAA 988
DB 142 CATATTAATTAAGTATTAACCTATTAATAATTTGAAAAATGGAATTTATTTCTTTTAA 201
QY 989 AACAA--TGATGCTAAATATTTTAAACACACAAATTAACCTTTTAAAGCA 1045
DB 202 AACAACTTCTATATAGAACTTTTAAACAAAAATGTATCATTTTAAACAAATTTGAAAGCG 261
QY 1046 ATCCTAATAGGAAACGAGGAAGTTAAAGATT 1076
DB 262 TGCTAACGAAAACAAGGAAGTTGAGAAAT 292

RESULT 7
ABL49954
ID ABL49954 standard; DNA; 424 BP.
XX

```

AC ABL49954;
 XX 10-JUN-2002 (first entry)
 DT
 XX Rice Wanderer-Osl nucleotide sequence SEQ ID NO:29.
 DE
 XX Rice; plant; detection; polymorphism; transposable element; genome;
 KW gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN WO200212484-A1.
 XX
 PD 14-FEB-2002.
 XX
 PF 02-AUG-2001; 2001WO-JP06661.
 XX
 PR 02-AUG-2000; 2000JP-0234577.
 XX
 PA (NISR) JAPAN TOBACCO INC.
 PA (SYGN) SYNGENTA LTD.
 XX
 PI Komori T, Nitta N;
 DR
 WP1; 2002-241759/29.
 XX
 CC Construction of marker for detecting plant genome polymorphism with use
 PT of transposable element, useful particularly in studying restriction
 PT fragment length polymorphism applicable in cell genetics -
 XX
 PS Example 2; Fig 7; 110pp; Japanese.
 XX
 CC The present invention describes a method for constructing a marker for
 CC detecting polymorphisms in a plant genome. The method comprises the
 CC production of a primer for nucleic acid amplification by using the base
 CC sequences of a transposable element and/or the domain adjacent to it.
 CC Also described is a marker for detection polymorphisms in plant genomes.
 CC The constructed marker can be used for detecting plant genome
 CC polymorphisms, which is useful particularly in studying restriction
 CC fragment length polymorphism applicable in cell genetics e.g. for
 CC analysing and selecting specific breeds of plants. The method is simple,
 CC easy, less time consuming and not so laborious, e.g. in the study of
 CC less frequently occurring polymorphism between various species to enable
 CC identification of the site and isolation of the required gene.
 CC ABL49956 to ABL50036 represent nucleotide sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 424 BP; 146 A; 66 C; 56 G; 156 T; 0 other;
 Query Match 6.8%; Score 97.4; DB 24; Length 424;
 Best Local Similarity 69.7%; Pred. No. 6.4e-11;
 Matches 147; Conservative 0; Mismatches 61; Indels 3; Gaps 1;
 QY 869 TTGATCCTCTAGTGGAGGTTGATTTTCAACGCGATGATAACGAGAAAGCTCATTAGCA 928
 DB 82 TTAATCTCCACATTTGGAGATAAATTTGGTTGCGACGTAAACGAGAAAGCTCATTAGCA 141
 QY 929 CATTTACTTAGATATTAT 988
 DB 142 CATATTAATTAAGTATTAACATTTATATATATATATATATATATATATATATATATAT 201
 QY 989 AACAATGATGATCAATATATTTTAAACACACACACACACACACACACACACACACACAC 1045
 DB 202 AACAACTTCTATAGAAACACTTTTCAACAAACAAACAAACAAACAAACAAACAAACAAAC 261
 QY 1046 ATCCCTAATAGGAACGAGGAAGTAAAGATT 1076
 DB 262 GTGCTAACGGAACAAAGGAAGTTGAAGAAAT 292
 RESULT 8
 ABL32289/c
 ID ABL32289 standard; DNA; 5875 BP.

XX ABL32289;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Human immune system associated gene SEQ ID NO: 262.
 DE
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianemic; cytostatic; neutropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX WPI; 2002-130909/17.
 DR
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 262; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 5875 BP; 1836 A; 27 C; 1078 G; 2934 T; 0 other;
 Query Match 4.6%; Score 66.6; DB 24; Length 5875;
 Best Local Similarity 48.0%; Pred. No. 0.00024;
 Matches 252; Conservative 0; Mismatches 269; Indels 4; Gaps 2;
 QY 534 AAACAAAATTAATATCATATAAAAGAGGTTGTTAATTACAAACCATGTTTCGTACTA 593
 DB 2227 AAAAAAATATAAATCCAAAAAACCACCTATTATTATATAAACTCTAAACATTATAAAA 2168
 QY 594 CAACTCTAATTTGTAATTTCTTATTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAA 653
 DB 2167 CAACTCTAATTTTAAATTTTAAATTTTAACTCTACCAAAATATAAAATTTTCTTTCAATACAT 2108
 QY 654 TAAACGTAGA-CGGCTAAGCCCAACCCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCAC 712
 DB 2107 TTAACAAATCTCGTCCATATACTAACACGAAATTAACACTCTTTTTCATCTTTTCTTT 2048
 QY 713 GAAATATGATGTTTATTATATATGATTTTTTTTAAATTAACCTTTTACATAAATTTCTTT 772
 DB 2047 TATAACAAAAATTTATTTTCTTAAATACAAATTTAAATTTTAAATTTTAAATTTTATTT 1988
 QY 773 AGGAAACATATCATTTAATGTTTGAAGAAACGTCGACATAGAAACAACTAAGAACGATGAG 832
 DB 1987 ACTTCCAAAAAATTCATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1928

XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 330; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX Query Match 4.4%; Score 64; DB 24; Length 9642;
Best Local Similarity 45.8%; Pred. No. 0.00091;
Matches 307; Conservative 0; Mismatches 350; Indels 13; Gaps 2;
QY 403 ATTGAATTTGAGTCCAACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATATT 462
DB 4113 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4054
QY 463 AATTTTCTTTCATGAAAGCAATTCACCGTTCGAGAAATATGCTGCTATAAATAAGTA 522
DB 4053 AATTAATTTTCATATACCAACTAATATACATATATATATATATATATATATATATATAT 3994
QY 523 GTCTAGTCGAGAAACAAATTAATATACATATAAAGAGAGTTGTTTAATTAACAACCAT 582
DB 3993 CTCCTATTAACTATCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3934
QY 583 GTTTCGTACTACACTTAATTTGTAATTTCTTATTCAGTCACAAAATTCCTCAATTCCA 642
DB 3933 CCT-----AAAAAATTAACCCATAAATTTACTAAATCACACAAATCTATATAA 3880
QY 643 ATTAAGAAAAATAAAGTAGCGGCTAAGCCACCCATCTAAGGCTAAGTTCGAGAGGTG 702
DB 3879 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3820
QY 703 AAGTAGCGACGAAATAATGATGGTTTATTAATATGATTTTTTTTAAATAACTTTTCACAT 762
DB 3819 AAAAAATTCCTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3760
QY 763 AATTTCTTTAGGAACATATCATTTAATGTTTGAAGAACGTCACATAGAAAACATAA 822
DB 3759 CACTACCTTAAACACCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3700
QY 823 GAACATGATGTTGGGAAACAAGAGAAACACAGCCTTAAGGCTTCTTGATCTCTAGTT 882
DB 3699 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3640
QY 883 GGAGGTTGATTTTCAACGCGATGATAACGAGAAAGCTCATAGCACATTATTACTTTAGA 942
DB 3639 ATAACAAATAATACATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3580
QY 943 TATTTATTAATTAACCTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1002
DB 3579 ACCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3527
QY 1003 AATATTATTTTAAACACACCAATTTAAACCTTTTAAAGAGCTCTTAATAGCAACGA 1062
DB 3526 AACTATTCTTTTAAACACATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3467
QY 1063 GGAAGTTAAA 1072

Db 3466 AAAAAATAAA 3457
RESULT 11
AAS46310/C
ID AAS46310 standard; DNA; 6095 BP.
XX AC AAS46310;
XX 18-DEC-2001 (first entry)
XX Tumour suppressor gene derived chemically modified sequence #32.
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
OS Homo sapiens.
XX WO200168912-A2.
XX 20-SEP-2001.
XX 15-MAR-2001; 2001WO-EP02955.
XX 15-MAR-2000; 2000DE-1013847.
XX 06-APR-2000; 2000DE-1019058.
XX 07-APR-2000; 2000DE-1019173.
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602752/68.
XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX Claim 1; SEQ ID No 32; 27pp; English.
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC complementary sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;

```
Query Match      4.4%; Score 63; DB 22; Length 6095;
Best Local Similarity 44.6%; Pred. No. 0.0013;
Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;

QY 461 TTAATTTTTCATGAAACGCAATCAACCGTTGCGAGAAATATGCTGTATATAAATAAG 520
DB 2571 TTAATTTTTCATGAAACGCAATCAACCGTTGCGAGAAATATGCTGTATATAAATAAG 520
QY 521 TAGTCTAGTGCAGAACCAATTAATATACATATAAAGAGAGGCTGTTAATTAACAAC 580
DB 2511 TAGTCTAGTGCAGAACCAATTAATATACATATAAAGAGAGGCTGTTAATTAACAAC 580
QY 581 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 640
DB 2451 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 640
QY 641 CAATTAAAGAAAAATAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 700
DB 2391 CAATTAAAGAAAAATAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 700
QY 701 TGAAGTACGACGCAAAATATGATGTTTATTAATATGATGTTTATTAATATGATGTT 760
DB 2331 TGAAGTACGACGCAAAATATGATGTTTATTAATATGATGTTTATTAATATGATGTT 760
QY 761 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 820
DB 2451 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 820
QY 821 AAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCCTTAAGGCTTCTTCTTCTTCT 880
DB 2211 AAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCCTTAAGGCTTCTTCTTCTTCT 880
QY 881 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 940
DB 2151 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 940
QY 941 GATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
DB 2092 GATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
QY 1001 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 1060
DB 2092 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 1060

RESULT 12
ABL70150/c
ID ABL70150 standard; DNA; 6095 BP.
XX AC ABL70150;
XX DT 01-JUL-2002 (first entry)
XX DE Chemically treated cell signalling DNA sequence complementary to#20.
XX KW Cell signalling; cytosine methylation; cell signalling disease;
XX KW cancer; tumour; cytosine; ds.
XX OS Unidentified.
XX PN WO200202807-A2.
XX PD 10-JAN-2002.
XX PF 29-JUN-2001; 2001WO-EP07471.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
```

(EPIG-) EPIGENOMICS AG.
Olek A, Piepenbrock C, Berlin K;
WPI; 2002-154758/20.

Nucleic acid, useful for diagnosis and therapy of diseases associated with cell signalling e.g. cancer, comprises chemically modified genomic sequences of genes associated with cell signalling -

Claim 1; SEQ ID NO 40; 24pp+sequence listing; English.

The invention relates to a nucleic acid comprising a sequence of at least 18 bases of a segment of chemically pretreated DNA of genes associated with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or PNA-oligoners for detecting cytosine methylations, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records ABL70111-ABL70626 represent chemically pre-treated genomic DNA's of genes associated with cell signalling.

Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;

Query Match 4.4%; Score 63; DB 24; Length 6095;
Best Local Similarity 44.6%; Pred. No. 0.0013;
Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;

QY 461 TTAATTTTTCATGAAACGCAATCAACCGTTGCGAGAAATATGCTGTATATAAATAAG 520
DB 2571 TTAATTTTTCATGAAACGCAATCAACCGTTGCGAGAAATATGCTGTATATAAATAAG 520
QY 521 TAGTCTAGTGCAGAACCAATTAATATACATATAAAGAGAGGTTGTTAATTAACAAC 580
DB 2511 TAGTCTAGTGCAGAACCAATTAATATACATATAAAGAGAGGTTGTTAATTAACAAC 580
QY 581 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 640
DB 2451 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 640
QY 641 CAATTAAAGAAAAATAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 700
DB 2391 CAATTAAAGAAAAATAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 700
QY 701 TGAAGTACGACGCAAAATATGATGTTTATTAATATGATGTTTATTAATATGATGTT 760
DB 2331 TGAAGTACGACGCAAAATATGATGTTTATTAATATGATGTTTATTAATATGATGTT 760
QY 761 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 820
DB 2451 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 820
QY 821 AAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCCTTAAGGCTTCTTCTTCTTCT 880
DB 2211 AAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCCTTAAGGCTTCTTCTTCTTCT 880
QY 881 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 940
DB 2151 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 940
QY 941 GATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
DB 2092 GATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
QY 1001 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 1060
DB 2092 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAAGTCTTCTTCTTCTTCT 1060

```

Db 2032 AAAAATAAATAAATATATAAAACCCCAAAAAACATAAAAAAATAAATAAATAA 1973
Qy 1061 GAGGAAGTTAAAGATTCAACCGAAGTGTTCGATATCAAAAAAT 1103
Db 1972 TAAACGATATAAATAATTTTAAAAAATAAATAAACAATATAT 1930

RESULT 13
ABL32361/c
ID ABL32361 standard; DNA; 6095 BP.
XX
AC ABL32361;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 334.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anti-naemic; cytosine methylation; antiasthmatic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineuritic; anti-arteriosclerotic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 334; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
XX Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;
XX

Query Match 4.4%; Score 63; DB 24; Length 6095;
Best Local Similarity 44.6%; Pred. No. 0.0013;
Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;

Qy 461 TTAATTTTTCATGAAACGAAATTCACCGTTCGAGAAATATGCTGCATAAATAAG 520
Db 2571 TTACTTTTAAATTCACATAACAAATAAATACTAGCTTTTAAATTCACAACTTAACCA 2512

Qy 521 TAGTCTAGTCGAGAAACAAATTAATATCATATAAAGAAAGAGTTGTTAATACAAACC 580
Db 2511 TAATATATTATAAATAAATAAATAAATTTATCATTTCAAAATAATAATTTTCAATCTCAAAA 2452

```

```

Qy 581 ATGTTTCGTACTACAACTCTAATTTGTAAATTCATTTTCAGTCACAAAATTCGAATTC 640
Db 2451 ACTTCCCTTAAAAAAATATTACAAACTCTTCCTCTCTTTTAAAAATACAAAAA 2392
Qy 641 CAATTAAAGAAAAATAAACGTCAGCGCTAAGCCCAACCCATCTAAGCGCTAAGTTCCGAGAGG 700
Db 2391 TAAAAAATAAATAAATTTTACACAAAAAATAAATAAATAAATAAATAAATAAATAA 2332
Qy 701 TGAAGTACGACGAAATAATATGATGGTTTATTAATATGATTTTAAAAATACTTTCAC 760
Db 2331 TAAACGTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2272
Qy 761 ATAAATTTCTTTAGGAAACATATATTTAATGTTTGAATAAATCGTCACATAAAGAAACT 820
Db 2271 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2212
Qy 821 AAGACGATGAGTTGGGAAACAAAGAGAGAAAAACACAGCGCTTAAAGGCTTCTTGATCCTCTAG 880
Db 2211 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2152
Qy 881 TTGAGGTTGATTTTCCAAACGATGATATAACGAGAAAGCTCATTAGCACATTTATTACTTA 940
Db 2151 AAAAAATTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2093
Qy 941 GATATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1000
Db 2092 TAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2033
Qy 1001 ATAAATTTTAAAAACACACCAATTTTAAACCTTTTAAAGAGCATCTTAATAGGAAC 1060
Db 2032 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1973
Qy 1061 GAGGAAGTTAAAGATTCAACCGAAGTGTTCGATATCAAAAAAT 1103
Db 1972 TAAACGATATAAATAATTTTAAAAAATAAATAAACAATATAT 1930

RESULT 14
ABL34475/c
ID ABL34475 standard; DNA; 6095 BP.
XX
AC ABL34475;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human metastasis associated gene SEQ ID NO: 28.
XX
KW Metastasis associated gene; cytostatic; gene therapy; cancer;
KW cytosine methylation; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200177376-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-EP03970.
XX
XX 06-APR-2000; 2000DE-1019058.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-010922/01.
XX
XX New nucleic acid derived from chemically treated metastasis genes,
XX useful for diagnosis of cancers by analysis of cytosine methylation,
XX also for treatment -
XX

```


Db 4082 CCATCTCAAAACAATTATTCACTTCACTTAAAAATTAAAAATATATCTCTAACCAATACA 4023
Qy 680 TCTAAGGCTAAGTTTCGAGAGTGA-AGTACGACGAGAAATATATGATGGTTTATTATATATG 738
Db 4022 TCTAAATTTATTATAATATCATTTTACTTAATAAAATATTTTCCAAATATTACCTA 3963
Qy 739 ATTTTAAATAAATCTTTCACATAAATTTCTTTAGGAAACATATCATTTTAAATGGTTGA 798
Db 3962 AATCTAATCGTCTAAACCCCTTTTCCAAATAAATAAAAAAATTTAAATTTCAATAATTATCA 3903
Qy 799 AAAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGGAAACAAGAGAAAAACAGCC 858
Db 3902 CTATTAAATTTTATCTAACCTTTAAACTTAATTAATTCATCCCACTATATTACACTAATAT 3843
Qy 859 TTAAGGCTTCTTGAPCCTCTAGTTGGAGGTTGATTTTCAAACGCATGATAAACGAGAAAG 918
Db 3842 ATAAAACTCCCAAAACGTAATAAATAAACTCATCTATCTTTACGAATAAATATCAAAATTA 3783
Qy 919 CTCATTAGCACATTTATTACTTAGATATTATAATATATAACTTGAAAAAATATTATTIT 978
Db 3782 ACTTACTTCATAAAATTAATATATTTTCATTAAATATTTTAAATAATTTTAAAAAATATATAAAA 3723
Qy 979 GAATTTTAAACAATGTATGCATAAATTTATTTTAAAAACACACCAATTTAACCCCTT 1038
Db 3722 ATATAAATTAATAATTTTAAAAAATAATTAATTTATACAATACACCTATTAAAAACCTT 3663
Qy 1039 AAAAGCATCCTTAATAGGAACGAGGAGTTAAAGATTAC 1079
Db 3662 TAAATTTATATTAAAAATTTAAAAAATAATTAACAATAAC 3622

Search completed: March 15, 2003, 10:56:21
Job time : 329.521 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:55 ; Search time 45,5838 Seconds
(without alignments)
9687,969 Million cell updates/sec

Title: US-09-802-937-2
Perfect score: 1440
Sequence: 1 ctcgaggaagattaattagg.....cgcgacgcaggtctgcgac 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| C 1 | 57.2 | 4.0 | 7218 | 1 US-08-232-463-14 | Sequence 14, Appl |
| C 2 | 50.2 | 3.5 | 6124 | 4 US-08-213-419B-3 | Sequence 3, Appl |
| C 3 | 50 | 3.5 | 19124 | 2 US-08-487-826B-13 | Sequence 13, Appl |
| C 4 | 46.8 | 3.2 | 636 | 4 US-08-998-416-1137 | Sequence 1137, Ap |
| C 5 | 46.6 | 3.2 | 1850 | 3 US-08-617-860B-32 | Sequence 22, Appl |
| C 6 | 46.6 | 3.2 | 2389 | 1 US-08-153-563-2 | Sequence 2, Appl |
| C 7 | 46.6 | 3.2 | 2389 | 2 US-08-460-507-2 | Sequence 2, Appl |
| C 8 | 46.6 | 3.2 | 4098 | 2 US-08-605-106-4 | Sequence 4, Appl |
| C 9 | 46.6 | 3.2 | 4276 | 1 US-07-973-324A-3 | Sequence 3, Appl |
| C 10 | 46.6 | 3.2 | 4276 | 4 US-08-343-380-3 | Sequence 3, Appl |
| C 11 | 46.6 | 3.2 | 4276 | 4 US-09-072-435-3 | Sequence 3, Appl |
| C 12 | 46.6 | 3.2 | 4276 | 4 US-09-072-917A-3 | Sequence 3, Appl |
| C 13 | 45.8 | 3.2 | 5852 | 1 US-07-867-106-2 | Sequence 2, Appl |
| C 14 | 45.8 | 3.2 | 9048 | 3 US-08-973-273-4 | Sequence 4, Appl |
| C 15 | 45 | 3.1 | 636 | 4 US-08-998-416-1137 | Sequence 1137, Ap |
| C 16 | 44.2 | 3.1 | 837 | 4 US-08-998-416-288 | Sequence 288, App |
| C 17 | 44 | 3.1 | 658 | 4 US-08-998-416-595 | Sequence 595, App |
| C 18 | 43.8 | 3.0 | 3950 | 4 US-09-398-395A-33 | Sequence 33, Appl |
| C 19 | 43 | 3.0 | 1857 | 4 US-09-299-378-3 | Sequence 3, Appl |
| C 20 | 42.8 | 3.0 | 6265 | 4 US-09-129-112-3 | Sequence 3, Appl |
| C 21 | 42.6 | 3.0 | 20674 | 4 US-09-641-638-651 | Sequence 651 App |
| C 22 | 42.6 | 3.0 | 28001 | 4 US-09-815-993-3 | Sequence 3, Appl |
| C 23 | 42.4 | 2.9 | 665 | 2 US-08-883-795A-36 | Sequence 36, Appl |
| C 24 | 42 | 2.9 | 711 | 4 US-08-998-416-786 | Sequence 786, App |
| C 25 | 41.8 | 2.9 | 3305 | 3 US-08-714-918-76 | Sequence 76, Appl |
| C 26 | 41.8 | 2.9 | 3305 | 4 US-09-265-315-76 | Sequence 76, Appl |
| C 27 | 41.8 | 2.9 | 3305 | 4 US-09-265-315-76 | Sequence 76, Appl |

| | | | | | |
|------|------|-----|-------|-----------------------|-------------------|
| C 28 | 41.8 | 2.9 | 3305 | 4 US-09-266-417-76 | Sequence 76, Appl |
| C 29 | 41.6 | 2.9 | 1431 | 4 US-09-316-083-2 | Sequence 2, Appl |
| C 30 | 41.6 | 2.9 | 6152 | 4 US-08-973-482-1 | Sequence 1, Appl |
| C 31 | 41.4 | 2.9 | 665 | 4 US-08-883-795A-36 | Sequence 36, Appl |
| C 32 | 41.4 | 2.9 | 2226 | 4 US-08-924-629C-35 | Sequence 35, Appl |
| C 33 | 41.2 | 2.9 | 1983 | 4 US-09-134-001C-2202 | Sequence 2202, Ap |
| C 34 | 41.2 | 2.9 | 5852 | 1 US-07-867-106-2 | Sequence 2, Appl |
| C 35 | 41.2 | 2.9 | 35060 | 3 US-08-814-095-7 | Sequence 7, Appl |
| C 36 | 41.2 | 2.9 | 84495 | 4 US-09-797-906-3 | Sequence 3, Appl |
| C 37 | 41 | 2.8 | 592 | 4 US-09-328-111-39 | Sequence 39, Appl |
| C 38 | 41 | 2.8 | 2061 | 3 US-09-020-033-1 | Sequence 1, Appl |
| C 39 | 40.8 | 2.8 | 724 | 4 US-08-998-416-683 | Sequence 683, App |
| C 40 | 40.2 | 2.8 | 3850 | 3 US-08-329-799-34 | Sequence 34, Appl |
| C 41 | 40.2 | 2.8 | 4285 | 4 US-09-410-464-1 | Sequence 1, Appl |
| C 42 | 40.2 | 2.8 | 6243 | 2 US-09-056-075-1 | Sequence 1, Appl |
| C 43 | 40 | 2.8 | 2058 | 2 US-08-749-391-1 | Sequence 1, Appl |
| C 44 | 40 | 2.8 | 2058 | 3 US-09-390-200-1 | Sequence 1, Appl |
| C 45 | 40 | 2.8 | 6124 | 4 US-08-213-419B-3 | Sequence 3, Appl |

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptz9pt-F1s
US-08-232-463-14
Query Match 4.0%; Score 57.2; DB 1; Length 7218;

Best Local Similarity 9.5%; Pred. No. 0.00011;
Matches 38; Conservative 197; Mismatches 165; Indels 0; Gaps 0;

```
QY 955 TAAACTTGAAAAAATATTTATTTGAAATTTTAAACAATATGATGCAATAAATATTTT 1014
DB 1515 TCAGTTTCAAAAAACGGCATGTAGGCATCACTGTAAATACCTATCTATGCAAGTAGTTAA 1456
QY 1015 AAAAAACACACCAATTTAAACCTTTTAAAGAGCATCTTAATAGNAACGAGGAAGTTAAAGA 1074
DB 1455 AGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1396
QY 1075 TTCACCAAGTGTGGATAATGAAAAATGGGTGGGATAGAAATGTTGTAATGAATCAG 1134
DB 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1336
QY 1135 GGTTAGGATTAATATTAATAATGAAGAGGAGAGATGAATGGTTAGAGTTTAAATGTGTC 1194
DB 1335 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1276
QY 1195 TTTTGGTGGTAGAAAAATATTTGCCATACACTCCCGAGAGCGGTGGTGGTGGT 1254
DB 1275 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1216
QY 1255 GGCAGAGCGCTCTTTTCCTTGGAAAAAATACTCTTAAAGGAAACAGAAAGAGGCC 1314
DB 1215 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1156
QY 1315 AGCTTTGGTGCACCTCTCACCAAGAAACAGAAACAAAAG 1354
DB 1155 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1116
```

RESULT 2

US-08-213-419B-3/C
Sequence 3, Application US/08213419B
Patent No. 6333406

GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.

TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
FILE OF INVENTION: AND USES THEREFOR

FILE REFERENCE: J11-002CNCIP

CURRENT APPLICATION NUMBER: US/08/213,419B

CURRENT FILING DATE: 1994-03-14

PRIOR APPLICATION NUMBER: US 07/870,506

PRIOR FILING DATE: 1992-04-17

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 3

LENGTH: 6124

TYPE: DNA

ORGANISM: Plasmodium falciparum

FEATURE:

NAME/KEY: CDS

LOCATION: (2407)..(2439)

NAME/KEY: CDS

LOCATION: (2598)..(3404)

NAME/KEY: CDS

LOCATION: (3580)..(3720)

NAME/KEY: CDS

LOCATION: (3850)..(5835)

US-08-213-419B-3

Query Match

Best Local Similarity 44.9%; Pred. No. 0.0053;

Matches 276; Conservative 0; Mismatches 333; Indels 6; Gaps 2;

```
QY 431 AATACGATTTTCTTTCAAAAGAAATTTATTAATTTTTTTCATGAACGCAATTCAA 490
```

```
DB 1188 AATAATTTATCTCTAGTACTTACCATATATATATTTTTTTTATAAGCAITTTAATA 1129
```

```
QY 491 CCGTTCGAGAAATATGCTGCTCAATAAAGTAGTCTAGTCGAGAAACAAATTAATCA 550
```

```
DB 1128 CATTTTATAGAGGTCTCTTAATAAAAAATTTGTTGGTAATACTACCCCAATCATTTT 1069
```

```
QY 551 CATAAAAAAGAGGTTGTTAAATTTACAAACCATGTTTCGTACTACAACTCTAATTTGTA 610
DB 1068 ATTATTTTCATA---TATATATGTATATATATATATATGCTAAATTTAAAGATGA 1013
QY 611 TTTCTATTTTCAGTCACAAAATTCCAATTTTCCAATTAAGAAAAAATAAAGCTAGACG 670
DB 1012 ATAACATATTAGTTATATAGCTATTACATATATGTACATTTTTCATTTCTCTCTTT 953
QY 671 GCCACCACTCTAAGGCTAAGTTTCGAGAGGTGAGCTACGCACGAAAAATATGATGTT 730
DB 952 ATATATTTAAAAAATAATATATATGTTTCATTAACCTCATAAATATATATATATATA 893
QY 731 TTAATATGATTTTTTTTTTAAATAAATCTTCA--CATAAATTTCTTTAGGAAACATAT 788
DB 892 TATATATATTTTTTTTTTTTAAATCTCATGAATACGCTTCCATTTTCGTAATGGAAT 833
QY 789 AATGGTTTGAAGACGTGCACATAAGAAAACCTAAGAACGATGAGTTGGGAAACAAGAA 848
DB 832 ATTTTTTCCCATATGATATATATATATATATATATATATTAAGACAAAATAAGAA 773
QY 849 AAACACAGCCTTAAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCA 908
DB 772 AAAAAGAAATACATAATATAGTTTAACTAAGAACTAGTTTAAATTCATATTTTAA 713
QY 909 AACGAGAAAGCTCATTAGCACATTTACTTTAGATATTTTATATATATATAAACTT 968
DB 712 TAAATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 653
QY 969 ATATTTATTTGAAATTTTTTAAACAATGTATGCATAAATTTATTTTTTAAACACAC 1028
DB 652 AACAAACTTGTATAAAATTTTCAATATGATTATTCATTTATATATATTTTCATA 593
QY 1029 TTAACCCCTTTAAAAA 1043
DB 592 TTTATACATAACAGA 578
```

RESULT 3

US-08-487-826B-13

Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobb Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826R-13

| Query Match | 3.5% | Score 50; | DB 2; | Length 19124; |
|-----------------------|---|-------------------|-----------------|-------------------|
| Best Local Similarity | 45.0%; | Fred. No. 0.0085; | | |
| Matches 273; | Conservative | 0; | Mismatches 325; | Indels 8; Gaps 2; |
| Qy 483 | CAATTCAACCGTTTCGAGAAATATGCTGTCATAAATAAGTAGTCTAGTCGACGAAACAAAAT | 542 | | |
| Db 15593 | CAAAAAAAATTAATAAAAAATTTTATATAAAAAAAATGATTATATAAAAAATATAAAA | 15652 | | |
| Qy 543 | TAATATCACATAAAAAAGAGGTCGTTTAATTACAAAAACATGTTTCGTAACCTACAACTCTAA | 602 | | |
| Db 15653 | CAAAAGAGAAAAAATAAAAAACATTAATAAAAAAATAATATATATATATAAAAAACAAAAA | 15712 | | |
| Qy 603 | TTTGTAAATTCCTATTTCAGTCACAAAAATCCAAATTTCCAAATTAAGAAAAATTAACGCTAG | 662 | | |
| Db 15713 | AAAGAAAAAAATATATTAAANATAAAAAATATATATATATATATATATATATATATATAT | 15771 | | |
| Qy 663 | ACGGCTAAGCCACCCATCTAAGGCTTAAGTTCGAGAGTGAAGTACGCAGAAAAATATG | 722 | | |
| Db 15772 | ATCTTAATAAAAAAATAATATACATAAAATAAAAAAATTTATTTAAATAAAAAAATAATA | 15831 | | |
| Qy 723 | ATGGTTTATTAATATGATTTTTTTTAAATAACCTTTTCACATAAAATTTCTTTAGGAACATA | 782 | | |
| Db 15832 | TAAATAAAAAAATTTAAATTTAAATAAAAAAATAATATAAAAAAATTTTAAATTAATAA | 15891 | | |
| Qy 783 | TCATTTAATGTTTGAATAACGTCACATAAGAAAACTTAAGAACGATGAGTTCGGAAAACA | 842 | | |
| Db 15892 | AAAAAAATTAATAAAAAATTTAATGAAATAAAAAAATAAAAAAATTTAAATTAATAAAA | 15951 | | |
| Qy 843 | AGAGAAAAACACAGCCTTAAGCCTTCTTGATCCTCTAGTTGGAGGTGTGATTTTCAAACGC | 902 | | |
| Db 15952 | AAATAAATAAATAATTAATTCATGCACATATACA-----TACATATATATATATATAT | 16004 | | |
| Qy 903 | ATGATTAACGAGAAGCTCATTTAGCACATTAATTACTTAGATATATTAATAATTAACACTTG | 962 | | |
| Db 16005 | ACCCATAAAGCTTACATACAACATTTTACACATACATATATATATATATATATATATATAT | 16064 | | |
| Qy 963 | AAAAAAATATTTATTTGAAATTTTTTAAAACAATGTATGCAATAAATTTATTTTTTAAAAAC | 1022 | | |
| Db 16065 | CATACACATTTTACACATACAT | 16124 | | |
| Qy 1023 | ACCAATTTAAACCTTTTAAAAAGCATCTTAATAGGAAGGAGGAAGTTTAAAGATTCACCGA | 1082 | | |
| Db 16125 | ATACATATATACATTAACAAACACATATATAATACCTAAATACATATATATACATACATA | 16184 | | |
| Qy 1083 | AGTGGT 1088 | | | |
| Db 16185 | TATGTT 16190 | | | |

```

RESULT 4
US-08-998-416-1137/C
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen
; APPLICANT: Knechtle, Philipp

```

```

1  APPLICANT:  Rebischung, Corinne
2  TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
3  TITLE OF INVENTION:  AND USES THEREOF
4  NUMBER OF SEQUENCES:  1152
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE:  No. 6239264artis Corporation
7  STREET:  3054 Cornwallis Road
8  CITY:  Research Triangle Park
9  STATE:  No. 6239264th Carolina
10 COUNTRY:  USA
11 ZIP:  27709
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  PatentIn Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/998,416
19 FILING DATE:  24-DEC-1997
20 CLASSIFICATION:  435
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:  CH 0016/97
23 FILING DATE:  31-DEC-1996
24 ATTORNEY/AGENT INFORMATION:
25 NAME:  Meigs, J. Timothy
26 REGISTRATION NUMBER:  38,241
27 REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE:  919-541-8587
30 TELEFAX:  919-541-8689
31 INFORMATION FOR SEQ ID NO:  1137:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH:  636 base pairs
34 TYPE:  nucleic acid
35 STRANDEDNESS:  single
36 TOPOLOGY:  linear
37 MOLECULE TYPE:  DNA (genomic)
38 ORGANISM:
39 ORGANISM:  PAG1692RP
40 J-08-998-416-1137

```

| Query Match | 3.2% | Score 46.8; | DB 4; | Length 636; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 44.6%; | Prod. No. 0.017; | | |
| Matches 274; | Conservative 0; | Mismatches 332; | Indels 8; | Gaps |
| Qy | 446 | TTCAAAGAAATATTAAATTTTTTTTCATGAAGCAAATTCACCGTTCGAGAAATAT | 505 | |
| Db | 636 | TTTTATAGTAGTATTTTAACTACATCTCTTATATATTTTATTAATTAATTAATGAT | 577 | |
| Qy | 506 | GCTGTCATAAATAAGTAGTCTAGTCGAGAAACAAATTAATATCACATAAAGAAAGGT | 565 | |
| Db | 576 | AAATATTAAATTAATAATTAATTAATTAATAATAATAATAATAATAATAATAAT | 517 | |
| Qy | 566 | TGTTAATTACAAACCGATGTTTCGTACTACACCTCAATTTTGTAATTTCTATTTCAGTCA | 625 | |
| Db | 516 | TATTAATAATATATAAAATTAATAATAAGAAATTAAGTTAAATTAATTTAATAATAA | 457 | |
| Qy | 626 | CAAAATTCCAATTTTCCAAATTAAGAAAAATAAAGCTAGAGCGGCTAAGCCACCCATCTAAG | 685 | |
| Db | 456 | TTCTTTAAAAAGAGTTAAATAATATAAAACAACATAATTTTATAAAAAATAGATATTATA | 397 | |
| Qy | 686 | GCTAAGTTCGAGAGGTGAAGTAGCGACGAAAAATATGATGGTTTATTA-----ATATG | 738 | |
| Db | 396 | ATAAAAATAAATAATTACAAATTTTAAATAAAAAATTAATCTTTTATAAATAAATAATATT | 337 | |
| Qy | 739 | ATTTTTTTTTAAATAACTTTCACATAAATTTCTTTAGGAAACATATCATTTTAATGGTTTGA | 798 | |
| Db | 336 | ATTTTAATAAACAAATTAATAAATAATATTAATTAATGATAACTATTTAATTAATTTAT | 277 | |
| Qy | 799 | AAACGTGCACATAGAAAACTAAGAACGATGAGTTGGGAAACAAAGAGAAAAACACAGCC | 858 | |
| Db | 276 | TAAAGAAAAATATATATCTCAATAATAATTTTATTAACACTAAITTTAAAAATTTGCAACATAGAC | 217 | |

```
QY 859 TTAAGGCTTCTTGATCCTCTAGTGGAGGTGATTTTCAAAAGCGATGATAAACGAGAAAG 918
| | | | |
Db 216 TAAATAGTATTCATATTAATAATATTTTATAATATATATAATATTAATATGATGAAT 157
| | | | |
QY 919 CTCATTAGCACATTTACTTACATATTTATATATTAATTAATTAACCTTGAAAAAATATTTTATTT 978
| | | | |
Db 156 TAAGTAATATATATATATATATATATATATATATATATATATATATATATATATATATTA 98
| | | | |
QY 979 GAATTTTAAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
| | | | |
Db 97 AATAATGATAATATAGTTTAAATATATTTAAATATATTTAAATATATTTAAATATATTTAAATATAT 38
| | | | |
QY 1039 AAAAAGCATCCTAA 1052
| | | | |
Db 37 ATTAATCTTTATAA 24
| | | | |

RESULT 5
US-08-617-860B-32/c
; Sequence 32, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Tyfofer, R., Bautor, J., Bothmann, H., Filsak, B.,
; APPLICANT: Hwike-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Miller, A., Schulte, W., Vostz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
; STREET: 1140 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,860B
; FILING DATE: 01-MAR-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02950
; FILING DATE: 05-SEP-1994
; APPLICATION NUMBER: DE P4329951.2
; FILING DATE: 04-SEP-1993
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1850 Base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: ClEg1
; FEATURE:
; NAME/KEY: CAAT-Signal
; LOCATION: 1428..1432
; FEATURE:
; NAME/KEY: TATA-Signal
; LOCATION: 1553..1556
; FEATURE:
; NAME/KEY: Transcription start
; LOCATION: 1585
; FEATURE:
; NAME/KEY: Leguminbox
```

```
; LOCATION: 1642..1657
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 1797..1799
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1797..1850
US-08-617-860B-32

Query Match      3.2%; Score 46.6; DB 3; Length 1850;
Best Local Similarity 47.2%; Pred. No. 0.026;
Matches 174; Conservative 0; Mismatches 194; Indels 1; Gaps 1;

QY 651 AATAAAACGCTAGACGGCTAAGCCACCCCATCTAAGGCTTAAGTTTCGAGAGGTGAAGTAGCG 710
| | | | |
Db 480 AATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 421
| | | | |
QY 711 ACGAAAAATATGATGGTTTATTAATATGATTTTAAATAAATCTTACATATAATTTCT 770
| | | | |
Db 420 AAAATTTTTCATATATTTCAAAAATTCAAAAATTCAAAAATATTTAAAAAATTTTAAATATATTTT 361
| | | | |
QY 771 TTAGGAACATATCATTTTAATGTTTGAAGAAACGTCACATAGAAAACTAAGAACGATG 830
| | | | |
Db 360 AAAACCAAAAAATATTTTAAAAATTCACAATATTTTAAAAATAAAAAACTAATTTTA 301
| | | | |
QY 831 AGTTGGGAAACAGAGAAACACACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTG 890
| | | | |
Db 300 AATATTTTAAAAAATTTTAAAAAATTAATTTTAAATATTTTAAAAAATTTTAAAAAATTTT 241
| | | | |
QY 891 ATTTTCAACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950
| | | | |
Db 240 AATTTAAAAATATTT-TAAATTTTAAAAAATATTTTAAAAAATTTTAAAAAATATTTTCAATA 182
| | | | |
QY 951 ATTATAACCTTGAAAAAATATTTTAAATTTTAAACCAATGTATGCATATAATTTAT 1010
| | | | |
Db 181 ATTCAATAATATTAACAATATATCCGCAAAATTTTAAATAGCGAATATTTCCCGAATA 122
| | | | |
QY 1011 TTTTAAAAA 1019
| | | | |
Db 121 TTCGTAATA 113
| | | | |

RESULT 6
US-08-153-563-2/c
; Sequence 2, Application US/08153563
; Patent No. 5693506
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond L.
; TITLE OF INVENTION: PROCESS FOR PROTEIN PRODUCTION
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,563
; FILING DATE: 16-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 2307E-515
; TELECOMMUNICATION INFORMATION:
; NAME/KEY: (415) 543-9600
```

TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2389
; OTHER INFORMATION: /standard_name="RAMY-1A"
US-08-153-563-2

Query Match 3.2%; Score 46.6; DB 1; Length 2389;
Best Local Similarity 51.7%; Pred. No. 0.029;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAGAAAACACAGCCTTAAGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAA 898
Db 2075 AACATGAATACCTGATCGCACAGAGCATTTGAGAGCAGAGAGAAAGAAATCGATTGAGA 2016

QY 899 ACCGATGATAAAGAGAGAAAGCTCAATTAGCACATTTACTTAGATATTTATAATTATAAA 958
Db 2015 ATGTACGCAAAAGAGCTGAGCCATTAAGTATATTAATTGAGTATTAACATAATTTTAA 1956

QY 959 CTTGAAAAAATATTTATTTGAATTTTAAACAATGTATGATCAATAATTTTAAAA 1018
Db 1955 TTTTAAAAATAGATAATAATAATTTTAAAGTAACTTTCCTATAGAAAAATTTTGCRAAAA 1896

QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
Db 1895 ATCATACCGTTTAGTAGTTTCAGGAA 1871

RESULT 7
US-08-460-507-2/c
; Sequence 2, Application US/08460507
; Patent No. 5994628
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,507
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,563
; FILING DATE: 16-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 2000-0452.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2389
; OTHER INFORMATION: /standard_name="RAMY-1A"
US-08-460-507-2

Query Match 3.2%; Score 46.6; DB 2; Length 2389;
Best Local Similarity 51.7%; Pred. No. 0.029;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAGAAAACACAGCCTTAAGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAA 898
Db 2075 AACATGAATACCTGATCGCACAGAGCATTTGAGAGCAGAGAGAAAGAAATCGATTGAGA 2016

QY 899 ACGATGATAAAGAGAGAAAGCTCAATTAGCACATTTACTTAGATATTTATAATTATAAA 958
Db 2015 ATGTACGCAAAAGAGCTGAGCCATTAAGTATATTAATTGAGTATTAACATAATTTTAA 1956

QY 959 CTTGAAAAAATATTTATTTGAATTTTAAACAATGTATGATCAATAATTTTAAAA 1018
Db 1955 TTTTAAAAATAGATAATAATAATTTTAAAGTAACTTTCCTATAGAAAAATTTTGCRAAAA 1896

QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
Db 1895 ATCATACCGTTTAGTAGTTTCAGGAA 1871

RESULT 8
US-08-605-106-4/c
; Sequence 4, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schweigman, Lundberg, Woessner & Kluch, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235.001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4098 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double stranded

```

,
, TOPOLOGY: linear
, MOLECULE TYPE: : DNS (genomic)
, HYPOTHETICAL: No
, ANTI-SENSE: No
, ORIGINAL SOURCE:
, ORGANISM: Cuphea lanceolata
, IMMEDIATE SOURCE:
, LIBRARY: Genomic Lambda FIX II
, CLONE: ClTeg1
, FEATURE:
, NAME/KEY: CDS
, LOCATION: join(1797..2294, 2655..3303, 3391..3459, 3460..3671)
, FEATURE:
, NAME/KEY: Startcodon
, LOCATION: 1797..1799
, FEATURE:
, NAME/KEY: exon II
, LOCATION: 1787..2294
, FEATURE:
, NAME/KEY: intron II
, LOCATION: 2295..2657
, FEATURE:
, NAME/KEY: exon III
, LOCATION: 2658..2791
, FEATURE:
, NAME/KEY: intron III
, LOCATION: 2792..2897
, FEATURE:
, NAME/KEY: exon IV
, LOCATION: 2898..3011
, FEATURE:
, NAME/KEY: intron IV
, LOCATION: 3012..3131
, FEATURE:
, NAME/KEY: exon V
, LOCATION: 3132..3303
, FEATURE:
, NAME/KEY: intron V
, LOCATION: 3304..3390
, FEATURE:
, NAME/KEY: exon VI
, LOCATION: 3391..3459
, FEATURE:
, NAME/KEY: intron VI
, LOCATION: 3460..3671
, FEATURE:
, NAME/KEY: exon VII
, LOCATION: 3672..3941
, FEATURE:
, NAME/KEY: Stopcodon
, LOCATION: 3942..3944
,
, US-08-605-106-4

```

| | | | | |
|-----------------------|--------------|------------------|-----------------|-------------------|
| Query Match | 3.2%; | Score 46.6; | DB 2; | Length 4098; |
| Best Local Similarity | 47.2%; | Pred. No. 0.034; | | |
| Matches 174; | Conservative | 0; | Mismatches 194; | Indels 1; Gaps 1; |

| | | | |
|----|-----|---|-----|
| QY | 651 | AAATAAACGTAGACGGCTAAGCCCAACCACATCTAAAGGCTTAAGTTCGAGAGGTGAAGTACGC | 710 |
| | | | |
| Dd | 480 | AATATATATATATATATATATATATATATATATATAAATTCTCGGTTATCTCCAAAC | 421 |
| | | | |
| QY | 711 | ACGAAAAATATGATGGTTTTATTAAATATGATTTTTTTTAAATAACTTTTCACATAAATTTCT | 770 |
| | | | |
| Dd | 420 | AAAATTTTTTTCATATTTCAAATAATTCAAAAAATATTAAAAACGTTTAAATAATATATTTTT | 361 |
| | | | |
| QY | 771 | TTAGGAAACATATCATTTAAATGGTTTGAAAAACGTGCACATAAGAAACCTAAGAACGATG | 830 |
| | | | |
| Dd | 360 | AAAACCAAAAAATATTTTAAAAAATTCACAATATTTTTAAAAATAAAAAAACTTAATTTTA | 301 |
| | | | |
| QY | 831 | AGTTGGAAACAAGAGAAAAACACACGCCTTAAGGCTTCTTGATCCTCTAGTTCGAGGTTG | 890 |
| | | | |
| Dd | 300 | AAATATTTTAAAAAAATTTTAAAAACCTAAATTTTATATATTTTAAAAATATTTTTTTTAAAAAG | 241 |
| | | | |

QY 891 ATTTTCAACGCGATGATAACGAGAAGCTCAATTAGCACATTATTACTTAGATATTTTATA 950

Db 240 AATTTAAATATATTTTAAATTTTAAAAATATTTTAAAAATTTTAAAAATATTTTCAATA 182

QY 951 ATTATAAACTTGAATAAATATTTATTTTGAATTTTAAAAACAATGATGCAATAATTTT 1010

Db 181 ATTCATATAATATACAAATATATATCCGCAAAATTTTAAATAGCGAAATATTTCCCGAATA 122

QY 1011 TTTTAAAAA 1019

Db 121 TTCGTAAAA 113

RESULT 9

US-07-973-324A-3/c

; Sequence 3, Application US/07973324A

; Patent No. 5460952

; GENERAL INFORMATION:

; APPLICANT: Yu, Su-May

; APPLICANT: Liu, Li-Fei

; TITLE OF INVENTION: Gene Expression System Comprising the

; TITLE OF INVENTION: Promoter Region of the Alpha-Amylase Genes

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/973,324A

; FILING DATE: 04-NOV-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Borun, Michael F.

; REGISTRATION NUMBER: 25447

; REFERENCE/DOCKET NUMBER: 31149

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300

; TELEFAX: 312/474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4276 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Rice (Oryzae sativa)

; STRAIN: CV. M202

; IMMEDIATE SOURCE:

; LIBRARY: (EMBL) genomic

; CLONE: '-Amy7-C

; FEATURE:

; NAME/KEY: CDS

; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704

; LOCATION: ..3952)

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704

; LOCATION: ..3952)

; PUBLICATION INFORMATION:

; AUTHORS: Yu et al., Su-May

; TITLE: Regulation of '-amylase-encoding gene expression

; TITLE: in germinating seeds and cultured cells of rice


```
; JOURNAL: Gene
; VOLUME: in press
US-07-973-324A-3

Query Match          3.2%; Score 46.6; DB 1; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.035;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTGTGAGGTTGATTTCAA 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2069 AACATGAATACCTGATCGCACACAGAGCAATTTGAGAGCAGAGAAAGAAATCGATTGAGA 2010
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 899 ACGCATGATTAACAGAAAGCTCATTAGCACATATTACTAGATATTATTAATTATATA 958
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2009 ATGTACGCAAAACGAGCTGAGCCATTACGTATAATTAATTCAGTATTAACTATTTAAA 1950
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 959 CTGGAAGAAAAATATTATTGAAATTTTTTAAACAAATGTATGCATAAATTTATTTTAAA 1018
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1949 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAAATTTTGCAGAAA 1890
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1019 ACACACCAATTTAACCTTTAAAAA 1043
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1889 ATCATACCGTTTAGTTACAGAA 1865
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-08-343-380-3/c
; Sequence 3, Application US/08343380
; Patent No. 5712112
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Liu, Li-Fei
; TITLE OF INVENTION: Gene Expression System Comprising the
; TITLE OF INVENTION: Promoter Region of the Alpha-Amylase Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,380
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/973,324
; FILING DATE: 04-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25447
; REFERENCE/DOCKET NUMBER: 31149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Rice (Oryzae sativa)
; STRAIN: CV. M202
; IMMEDIATE SOURCE:

; LIBRARY: (EMBL) genomic
; CLONE: '-Amy7-C
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: Join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
; PUBLICATION INFORMATION:
; AUTHORS: Yu et al., Su-May
; TITLE: Regulation of '-amylase-encoding gene expression
; TITLE: in germinating seeds and cultured cells of rice
; JOURNAL: Gene
; VOLUME: in press
US-08-343-380-3

Query Match          3.2%; Score 46.6; DB 1; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.035;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTGTGAGGTTGATTTCAA 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2069 AACATGAATACCTGATCGCACACAGAGCAATTTGAGAGCAGAGAAAGAAATCGATTGAGA 2010
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 899 ACGCATGATTAACAGAAAGCTCATTAGCACATATTACTAGATATTATTAATTATATA 958
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2009 ATGTACGCAAAACGAGCTGAGCCATTACGTATAATTAATTCAGTATTAACTATTTAAA 1950
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 959 CTGGAAGAAAAATATTATTGAAATTTTTTAAACAAATGTATGCATAAATTTATTTTAAA 1018
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1949 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAAATTTTGCAGAAA 1890
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1019 ACACACCAATTTAACCTTTAAAAA 1043
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1889 ATCATACCGTTTAGTTACAGAA 1865
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-09-072-435-3/c
; Sequence 3, Application US/09072435
; Patent No. 6215051
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Liu, Li-Fei
; APPLICANT: Chan, Ming-Tsair
; TITLE OF INVENTION: GENE EXPRESSION SYSTEM COMPRISING THE
; TITLE OF INVENTION: PROMOTER REGION OF THE ALPHA-AMYLASE GENES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,435
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,792
; FILING DATE: 29-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/973,324
; FILING DATE: 04-NOV-1992
; ATTORNEY/AGENT INFORMATION:
```

NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28123/34274
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4276 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Rice (Oryza sativa)
STRAIN: CV. M202
IMMEDIATE SOURCE:
LIBRARY: (EMBL) genomic
CLONE: -Amy7-C
FEATURE:
NAME/KEY: CDS
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
LOCATION: ..3952)
FEATURE:
NAME/KEY: mat peptide
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
LOCATION: ..3952)
PUBLICATION INFORMATION:
AUTHORS: Yu et al., Su-May
TITLE: Regulation of -amylase-encoding gene expression
TITLE: in germinating seeds and cultured cells of rice
JOURNAL: Gene
VOLUME: in press
US-09-072-435-3

Query Match 3.2%; Score 46.6; DB 4; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.035;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 839 AACAGAGAAAAACACAGCCTTAAGCCTTCTAGTCTCTAGTGGAGGTTGATTTCAA 898
DB 2069 AACATGAATACCTGATCGCACACAGCATTGTGAGCAGAGAAAGAAATCGATTGAGA 2010
QY 899 ACGCATGATAAAGCAGAAAGCTCATTACACATTATTACTTAGATATTATTAATAA 958
DB 2009 ATGTAGCAAAACGAGCTGAGCCATTACGTATTAATTAATGAGTATTAATACTTTTAA 1950
QY 959 CTTGAAAAAATATTATTGTAATTTTAAACAATGTATGCATAAATTTATTTTAAAA 1018
DB 1949 TTTTAAAAATAGATTAAATATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCACAAA 1890
QY 1019 ACACACCAATTTAACCCCTTTAAAA 1043
DB 1889 ATCATACCGTTTAGTAGTTTCAGGAA 1865

RESULT 12
US-09-072-917A-3/c
Sequence 3, Application US/09072917A
Patent No. 6288302
GENERAL INFORMATION:
APPLICANT: Yu, Su-May
APPLICANT: Liu, Li-Fei
APPLICANT: Chan, Ming-Tsair
TITLE OF INVENTION: Application of Alpha-Amylase Gene
TITLE OF INVENTION: Promoter and Signal Sequence in the Production of
Patent No. 6288302
TITLE OF INVENTION: Recombinant Proteins in Transgenic Plants and Transgenic
TITLE OF INVENTION: Plant Seeds
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower

CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,917A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/509,962
FILING DATE: 01-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28123/34257
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4276 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Rice (Oryza sativa)
STRAIN: CV. M202
IMMEDIATE SOURCE:
LIBRARY: (EMBL) genomic
CLONE: alpha-Amy7-C
FEATURE:
NAME/KEY: CDS
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
LOCATION: ..3952)
FEATURE:
NAME/KEY: mat peptide
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
LOCATION: ..3952)
PUBLICATION INFORMATION:
AUTHORS: Yu et al., Su-May
TITLE: Regulation of alpha-amylase-encoding gene expression
TITLE: in germinating seeds and cultured cells of rice
JOURNAL: Gene
VOLUME: in press
US-09-072-917A-3
Query Match 3.2%; Score 46.6; DB 4; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.035;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 839 AACAGAGAAAAACACAGCCTTAAGCCTTCTAGTCTCTAGTGGAGGTTGATTTCAA 898
DB 2069 AACATGAATACCTGATCGCACACAGCATTGTGAGCAGAGAAAGAAATCGATTGAGA 2010
QY 899 ACGCATGATAAAGCAGAAAGCTCATTACACATTATTACTTAGATATTATTAATAA 958
DB 2009 ATGTAGCAAAACGAGCTGAGCCATTACGTATTAATTAATGAGTATTAATACTTTTAA 1950
QY 959 CTTGAAAAAATATTATTGTAATTTTAAACAATGTATGCATAAATTTATTTTAAAA 1018
DB 1949 TTTTAAAAATAGATTAAATATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCACAAA 1890
QY 1019 ACACACCAATTTAACCCCTTTAAAA 1043
DB 1889 ATCATACCGTTTAGTAGTTTCAGGAA 1865

[illegible]

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 15:53:40 ; Search time 97.3949 Seconds
(without alignments)
10377.192 Million cell updates/sec

Title: US-09-802-937-2

Perfect score: 1440
Sequence: 1 cgcgaggaagattaattagg.....cgcgacgcgagtgctgcac 1440

Scoring table: IDENTITY NUC
Gapop 10.0 , Capext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 1384.6 | 96.2 | 1443 | 10 | US-09-802-927-2 |
| 2 | 1384.6 | 96.2 | 4354 | 10 | US-09-802-927-1 |
| 3 | 62.4 | 4.3 | 2000 | 10 | US-09-887-576-828 |
| 4 | 55.8 | 3.9 | 8416 | 8 | US-08-910-386A-4 |
| 5 | 55.8 | 3.9 | 13341 | 8 | US-08-910-386A-1 |
| 6 | 55.8 | 3.9 | 19639 | 8 | US-08-910-386A-6 |
| 7 | 53.2 | 3.7 | 5940 | 8 | US-08-910-386A-10 |
| 8 | 50.4 | 3.5 | 2000 | 10 | US-09-887-576-843 |
| 9 | 50.4 | 3.5 | 335913 | 9 | US-09-754-853A-2 |
| 10 | 50.4 | 3.5 | 335913 | 9 | US-09-754-853A-3 |
| 11 | 50.2 | 3.5 | 33675 | 9 | US-09-921-992-2 |
| 12 | 48.8 | 3.4 | 335913 | 9 | US-09-754-853A-2 |
| 13 | 48.8 | 3.4 | 335913 | 9 | US-09-754-853A-3 |
| 14 | 47.2 | 3.3 | 840881 | 10 | US-09-790-988-1 |
| 15 | 46.6 | 3.2 | 419 | 10 | US-09-960-352-11234 |
| 16 | 45.4 | 3.2 | 2000 | 9 | US-09-938-842A-4747 |
| 17 | 44.6 | 3.1 | 442 | 10 | US-09-960-352-12911 |
| 18 | 44.6 | 3.1 | 2267 | 10 | US-09-822-830A-227 |
| 19 | 44.6 | 3.1 | 19639 | 8 | US-08-910-386A-6 |

| | | | | | | |
|----|------|-----|--------|----|---------------------|---------------------|
| 20 | 44.4 | 3.1 | 2000 | 9 | US-09-938-842A-3659 | Sequence 3659, Ap |
| 21 | 44.4 | 3.1 | 7204 | 8 | US-08-910-386A-11 | Sequence 11, Appl |
| 22 | 44.2 | 3.1 | 516 | 10 | US-09-960-352-5785 | Sequence 5785, Ap |
| 23 | 44.2 | 3.1 | 2000 | 9 | US-09-938-842A-4774 | Sequence 4774, Ap |
| 24 | 44 | 3.1 | 2000 | 9 | US-09-938-842A-3672 | Sequence 3672, Ap |
| 25 | 43.8 | 3.0 | 393 | 10 | US-09-960-352-4582 | Sequence 4582, Ap |
| 26 | 43.8 | 3.0 | 3950 | 10 | US-09-887-586A-33 | Sequence 33, Appl |
| 27 | 43.8 | 3.0 | 3950 | 10 | US-09-903-012-33 | Sequence 33, Appl |
| 28 | 43.6 | 3.0 | 449 | 10 | US-09-960-352-2412 | Sequence 2412, Appl |
| 29 | 43.2 | 3.0 | 277 | 10 | US-09-960-352-12673 | Sequence 12673, A |
| 30 | 43.2 | 3.0 | 342 | 10 | US-09-960-352-2399 | Sequence 2399, Ap |
| 31 | 42.8 | 3.0 | 1943 | 10 | US-09-960-731-2 | Sequence 2, Appli |
| 32 | 42.8 | 3.0 | 3759 | 12 | US-10-051-952-7 | Sequence 7, Appli |
| 33 | 42.8 | 3.0 | 6265 | 10 | US-09-129-112-3 | Sequence 3, Appli |
| 34 | 42.8 | 3.0 | 513509 | 9 | US-09-754-853A-4 | Sequence 4, Appli |
| 35 | 42.6 | 3.0 | 1152 | 9 | US-10-098-841-163 | Sequence 163, App |
| 36 | 42.6 | 3.0 | 28001 | 9 | US-10-193-295-3 | Sequence 3, Appli |
| 37 | 42.4 | 2.9 | 2000 | 9 | US-09-938-842A-2830 | Sequence 2830, Ap |
| 38 | 42.4 | 2.9 | 3056 | 9 | US-10-037-598-30 | Sequence 30, Appl |
| 39 | 42.4 | 2.9 | 3142 | 9 | US-10-037-598-37 | Sequence 37, Appl |
| 40 | 42.4 | 2.9 | 3410 | 9 | US-10-196-063-1 | Sequence 1, Appli |
| 41 | 42.4 | 2.9 | 74586 | 10 | US-09-781-558-3 | Sequence 3, Appli |
| 42 | 42.2 | 2.9 | 684973 | 10 | US-09-263-959-1 | Sequence 1, Appli |
| 43 | 42 | 2.9 | 513509 | 9 | US-09-754-853A-4 | Sequence 4, Appli |
| 44 | 41.8 | 2.9 | 431 | 10 | US-09-960-352-5558 | Sequence 5558, Ap |
| 45 | 41.6 | 2.9 | 446 | 10 | US-09-960-352-3400 | Sequence 3400, Ap |

ALIGNMENTS

RESULT 1
US-09-802-927-2
; Sequence 2, Application US/09802927
; Patent No. US20010031236A1
; GENERAL INFORMATION:
; APPLICANT: UCHIMIYA, HIROFUMI
; APPLICANT: ARAI, SATOSHI
; APPLICANT: FUSHIMI, TAKAOMI
; APPLICANT: TAGAWA, MICHITO
; APPLICANT: FUKUZAWA, HIROMITSU
; TITLE OF INVENTION: DNA FRAGMENT HAVING PROMOTER FUNCTION
; FILE REFERENCE: 204323USCIP
; CURRENT APPLICATION NUMBER: US/09/802,927
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: PCT/JF98/04088
; PRIOR FILING DATE: 1988-09-10
; PRIOR APPLICATION NUMBER: PCT/JF99/04847
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Oryza sativa L.cv.Nipponbare
US-09-802-927-2

Query Match 96.2%; Score 1384.6; DB 10; Length 1443;
Best Local Similarity 99.2%; Pred. No. 3e-265;
Matches 1434; Conservative 0; Mismatches 4; Indels 7; Gaps 4;

| | | | |
|----|-----|--|-----|
| QY | 1 | CTGCAGGAAGATTAATTAGTGGACACACCAACCCCTGTGTTGGTGGACCCCTGTTGTT | 60 |
| Db | 1 | CTGCAGGAAGATTAATTAGTGGACACACCAACCCCTGTGTTGGTGGACCCCTGTTGTT | 60 |
| QY | 61 | ATCACTGGGGTGTTCGTTGGACATGTTTTCAGAGAAATTAAGCAAGAAATTAAG | 120 |
| Db | 61 | AATCAACTGGGGTGTTCGTTGGACATGTTTTCAGAGAAATTAAGCAAGAAATTAAG | 120 |
| QY | 121 | AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTT | 180 |
| Db | 121 | AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTT | 180 |


```
QY 420 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATATTAATTTTTCATGAA 479
Db 421 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATATTAATTTTTCATGAA 480
QY 480 AGCAATTTCAACCGTTGAGAAATATGCTGCTATAAATAAGTAGTCTAGTGCAGAAACAA 539
Db 481 AGCAATTTCAACCGTTGAGAAATATGCTGCTATAAATAAGTAGTCTAGTGCAGAAACAA 540
QY 540 AATTAATATCATATAAAAGAGGTTGTTAATTAACAACCATGTTTGTACTACAATC 599
Db 541 AATTAATATCATATAAAAGAGGTTGTTAATTAACAACCATGTTTGTACTACAATC 600
QY 600 TAAATTTGTAATTTCTTATTTTCAGTCAAAAATTTCCAATTTCCAATTTCAAAAATTAACG 659
Db 601 TAAATTTGTAATTTCTTATTTTCAGTCAAAAATTTCCAATTTCCAATTTCAAAAATTAACG 660
QY 660 TAGAGCGCTAAGCCACCCATCTAAGCTAAGTTCGAGAGGTGAAGTAGCGACGAAAT 719
Db 661 TAGAGCGCTAAGCCACCCATCTAAGCTAAGTTCGAGAGGTGAAGTAGCGACGAAAT 720
QY 720 ATGATGGTTTAAATATGATTTTAAATTAACCTTTTACATAAATTTCTTTAGGAAC 779
Db 721 ATGATGGTTTAAATATGATTTTAAATTAACCTTTTACATAAATTTCTTTAGGAAC 780
QY 780 ATATCATTTAATGTTTGAAGAGTGCACATAAGAAACTAAGAACGATGATTTGGAA 839
Db 781 ATATCATTTAATGTTTGAAGAGTGCACATA--GAAACTAGAACGATGATTTGGAA 838
QY 840 ACAAGAGAAAACACAGCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAA 899
Db 839 ACAAGAGAAAACACAGCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAA 898
QY 900 CGCATGATAACGAGAGGCTATAGCATTATTAATTAATTAATTAATTAATTAATTAAC 959
Db 899 CGCATGATAACGAGAGGCTATAGCATTATTAATTAATTAATTAATTAATTAATTAAC 958
QY 960 TTGAAAAAATATTTATTTGAATTTTAAACAATGTATGCAATAATTTTAAAAA 1019
Db 959 TTGAAAAAATATTTATTTGAATTTTAAACAATGTATGCAATAATTTTAAAAA 1018
QY 1020 CACACCAATTTAACCTTTAAAAAGCATCTAATAGAAACGAGGAAGTTAAAGATTCCAC 1079
Db 1019 CACACCAATTTAACCTTTAAAAAGCATCTAATAGAAACGAGGAAGTTAAAGATTCCAC 1078
QY 1080 CGAA---GTGTTTGGATAATGAAATATGGGTGGGATTTAGAAATTTGTAATGAATCAGG 1136
Db 1079 CGAAGGTGTGTTTGGATAATGAAATATGGGTGGGATTTAGAAATTTGTAATGAATCAGG 1138
QY 1137 TTAGGATTAATATTTAAATGAAAGAGGAGAAATGAATGTTTAGAGTTTAAATGTGCTT 1196
Db 1139 TTAGGATTAATATTTAAATGAAAGAGGAGAAATGAATGTTTAGAGTTTAAATGTGCTT 1198
QY 1197 TTTGTTGGGTAGAAAATTTATTTGGCATACACTCCCGAGAGCGGTGCGTGTGGTGGG 1256
Db 1199 TTTGTTGGGTAGAAAATTTATTTGGCATACACTCCCGAGAGCGGTGCGTGTGGTGGG 1258
QY 1257 CAGAAAGCGTCTTTTTCGTTGGAAAAAATACTGCTTAAAAAGGAAACAGAAAGAGCCAG 1316
Db 1259 CAGAAAGCGTCTTTTTCGTTGGAAAAAATACTGCTTAAAAAGGAAACAGAAAGAGCCAG 1318
QY 1317 CTTT--GTTGTACCGTCTCACCAGAAACGAAACAAAAGCCCAACCACTTAACTCTC 1375
Db 1319 CTTTGGTGTACCGTCTCACCAGAAACGAAACAAAAGCCCAACCACTTAACTCTC 1378
QY 1376 GATCGACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGAGAGACGCGACGCGAGTTTCG 1435
Db 1379 GATCGACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGAGAGACGCGACGCGAGTTTCG 1438
QY 1436 TCGAC 1440
Db 1439 TCGAC 1443
```

```
RESULT 3
US-09-887-576-828/c
; Sequence 828, Application US/09887576
; Patent No. US2002014047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 828
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-828

Query Match 4.3%; Score 62.4; DB 10; Length 2000;
Best Local Similarity 61.9%; Pred. No. 0.0017;
Matches 99; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 903 ATGATAACGAGAAAGCTCATTTAGCACATTTACTTAGATTTTATTAATTAACCTTG 962
Db 1227 ACGTAAACGAGATGAGCCCATTAACGTATGATTAATTAAGTATTAACCTTTTAACTTT 1168
QY 963 AAAAAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATTTATTTTAAAAACAC 1022
Db 1167 AAAAAATAGATTATATGATTTTTTAAAGCACTTTCTTATAGAAATTTTTCGCAAAAAAC 1108
QY 1023 ACCAATTTAACCTTTAAAGCATCTTAATAGGAACGA 1062
Db 1107 ACACCGTTTAGTTCGAAACGTCGCGTGGAATGA 1068

RESULT 4
US-08-910-386A-4/c
; Sequence 4, Application US/08910386A
; Patent No. US2002092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(4771..7384, 7676..8052)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member A1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7432..7614
; OTHER INFORMATION: /note= "Snap-O11, transposon-like"
; OTHER INFORMATION: element"
;
; US-08-910-386A-4
```

Query Match 3.9%; Score 55.8; DB 8; Length 8416;
Best Local Similarity 57.0%; Pred. No. 0.05;
Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy 891 ATTTCACGCGATGATAACGAGAAAGCTCATTTAGCACATTATTACCTAGATATTATTA 950
Db 4246 AATTGACACATAGCTAGACGAAATAGCTGTAGTGTAATTAATAGTAGTAATT 4187
Qy 951 ATTATTAACCTGAAAAATATTTATTTGAATTTTAAACAAATGATATGCATAAATTATT 1010
Db 4186 ATTTAAACTTAAAAAATAGATATATATTTTAAAGAAACTTTTCTATATAAAGTTT 4127
Qy 1011 TTTTAAAAACACACCAATTTAACCCCTTTAAAAAGCATCTAATAGGAAACGAGGAAGTT 1069
Db 4126 TTTAAAAAACACACCATTTACAGTTCTGAAACATACGCGTGAACAAACGATCCTTT 4068
```

RESULT 5 US-08-910-386A-1/c

```
; Sequence 1, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scott
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2367..4205
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member D"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4201..9071
; OTHER INFORMATION: /note= "retrofit, a copia-like,
; OTHER INFORMATION: transposon-like element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4484..8821
; OTHER INFORMATION: /product= "retrofit"
; OTHER INFORMATION: /gene= "gag/pol"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 9915..11712
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10020..10975
; OTHER INFORMATION: /note= "Krispie, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 12626..12750
; OTHER INFORMATION: /note= "Pop-O12, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice2, transposon-like
; OTHER INFORMATION: element"
;
; US-08-910-386A-1
```

Query Match 3.9%; Score 55.8; DB 8; Length 13341;
Best Local Similarity 57.0%; Pred. No. 0.05;
Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy 891 ATTTTCAACGCGATGATAACGAGAAAGCTCATTTAGCACATTATTACCTAGATATTATTA 950
Db 1842 AATTGACACATAGCTAGACGAAATAGCTGTAGTGTAATTAATAGTAGTAATT 1783
Qy 951 ATTATTAACCTGAAAAAATATTTATTTGAATTTTAAACAAATGATATGCATAAATTATT 1010
```


Db 1782 ATTTAACTTAAAAATAGATTATATATTTTAAAGAACTTTTCTATAGAAAGTTT 1723
QY 1011 TTTTAAAAACACCACTTTTAAACCTTTTAAAGCACTCTTAATAGGAAACGAGGAAGTT 1069
Db 1722 TTTCAAAAACACACCACTTTTAAACGTTCTGAAACATACCGGTGAACACGATGCTTTT 1664

RESULT 6

US-08-910-386A-6/c
; Sequence 6, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: -
; LOCATION: 5213..18201
; OTHER INFORMATION: /note= "Xa21 gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(5213..7889, 8732..9132)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 disease resistance gene"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 9645..9769
; OTHER INFORMATION: /note= "Pop-Oil1, transposon-like"
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice1, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(15118..17720, 17827..18204)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member C; 2 bp
; OTHER INFORMATION: deletion causing frame-shift mutation of
; OTHER INFORMATION: ORF compared to family member A1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16183..16184
; OTHER INFORMATION: /note= "location of 2 bp deletion
; OTHER INFORMATION: compared to family member A1"
; US-08-910-386A-6
Query Match 3.9%; Score 55.8; DB 8; Length 19639;
Best Local Similarity 57.0%; Pred. No. 0.061;
Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 891 ATTTTCAAAACGATGATAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATA 950
Db 4698 AATTGACAAGATACGTAGAACGAAATAAGCTGTAGTGATAATTTAATTAAGTAGTAATT 4629
QY 951 ATTATAACTTGAAGAAAATATTATTGTAATTTTAAACAACTGATGCATAAATTATT 1010
Db 4628 ATTTTAAACTTTAAAAAATAGATTAATATATTTTAAAGAAACITTTTCTATAGAAAGTTT 4569
QY 1011 TTTTAAAAACACCACTTTTAAACCTTTTAAAGCACTCTTAATAGGAAACGAGGAAGTT 1069
Db 4568 TTTCAAAAACACACACCACTTTTAAACGTTCTGAAACATACCGGTGAACACGATGCTTTT 4510

RESULT 7

US-08-910-386A-10/c
; Sequence 10, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5940 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2151..5855
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "k21 family member A2; truncated
; OTHER INFORMATION: by two mutations"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(2501..2503, "acc")
; OTHER INFORMATION: /note= "mutation compared to family
; OTHER INFORMATION: member A1"
; FEATURE:
; NAME/KEY: mutation
; LOCATION: replace(4355..4356, "ac")
; OTHER INFORMATION: /note= "mutation compared to family
; OTHER INFORMATION: member A1"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5453..5697
; OTHER INFORMATION: /note= "Tourist-Ol2, transposon-like
; OTHER INFORMATION: element"
US-08-910-386A-10

Query Match
Best Local Similarity 3.7%; Score 53.2; DB 8; Length 5940;
Matches 139; Conservative 0; Mismatches 123; Indels 2; Gaps 1;

QY 908 AAACGAGAAAGCTCATTAGCAGCATATTACTTAGATATTTTAATAATTAACTTGAAGAAA 967
DB 5647 AAACGAGTGAGCCATTAGTACATGATTAAATGAGTATTAACATTTTAACTTCAAAA 5588

QY 968 AATATTTTGAATTTTAAACAA--TGTATGCATAAATTTTAAACACACACC 1025
DB 5587 TGGATTATATGATTTTAAAGCAACTTTCATATAGAAAATTTTTCGCAAAAACACAC 5528

QY 1026 AATTTAACCCCTTAAAGCATCTTAATAGGAACGAGGAGTTAAAGATTCCCGAAGT 1085
DB 5527 CGTTTAGTAGTTCGGAGAGCGCGGTGGAACAAACAACTACTCAATCTCCT 5468

QY 1086 GTTTGGATAATGAAAAATGGGGTGGGATTAGAATTTGTTAAATGAATCAGGGTTAGGATTA 1145
DB 5467 CTGCCGAACCCCTCCCTTAGGTGTGTGTTATTATGAGAAATACAAATAGAACAGATAC 5408

QY 1146 AATATTAATGAAGAGGAGGAGAA 1169
DB 5407 ATGATCTGATGGCAGATCAGGAA 5384

RESULT 8
US-09-887-576-843/c
; Sequence 843, Application US/09887576
; Patent No. US2002014047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576

; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 843
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-843

Query Match
Best Local Similarity 3.5%; Score 50.4; DB 10; Length 2000;
Matches 166; Conservative 0; Mismatches 111; Indels 19; Gaps 3;

QY 714 AAAAATATGATGGTTTATTAAATATGATTTTTTTTAAATAACTTTTCACATAAATTTCTTTA 773
DB 411 AAAACCTTTTACATATAAATTAAATATTATTGTTTAAACAAATTCCTTCATATTTTTTTT 352

QY 774 GGAAACATATCATTTTAATGTTTGAATAACGTCACATAAGAAACCTAAGAACGATGAGT 833
DB 351 AAAAATATACGTTTAGCTGTTTGA-----GCGTGGCCGGAATAAAGAGATAAG 297

QY 834 TGGGAAACAAGAGAAACACACAGCCCTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATT 893
DB 296 TTAATGAACAATGAATCCCGAACATCGAA-----CCTTGGGTCGTAACTAAT 250

QY 894 TTCAACGCGCATGATAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATTAATT 953
DB 249 TCCTCCCGCAGCAAAACAAATGTCGCTCATTAGCGCATGAGTAATTAAGTA-TTATCCTA 191

QY 954 ATAACCTTGAAAAATATTTATTGCAATTTTAAACAATGTATGCATAAATTAT 1009
DB 190 AAAACCTGAAAAAATGTTATTATTTCGATTTTTTAAAAAATATATAGATTCTTTTT 135

RESULT 9
US-09-754-853A-2
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; APPLICANT:
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45163)..(45314),(45450)..(45509)..(46941)..(48763)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-2

Query Match
Best Local Similarity 3.5%; Score 50.4; DB 9; Length 335913;
Matches 377; Conservative 0; Mismatches 436; Indels 15; Gaps 5;

QY 364 ACAAAATCGCAAAATATCATGGTTATCTATCTTCTGCTCAAAATGAAATTTAGTCCAACGT 423
```

| | | | |
|----|-------|--|-------|
| Db | 98994 | ATATATTTTGGTATTTAAATTTTATATAAATTTCTTAATTAAGTTATGTGAAGTTGATTTG | 99055 |
| Qy | 424 | AGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTCTTCATGAAACGC | 483 |
| Db | 99054 | TTTTGTAAATAAAATAAAAATTAACACAATGATATTAATATGTGTTTTCTTTTAA--A | 99111 |
| Qy | 484 | AAATTCAACGGTTCGAGAAATATGCTGTCATATAAATAAGTAGTCTAGTCGCAAAACAAATTT | 543 |
| Db | 99112 | TACTTAAATTTTAAAAACTAAAAATTTTAAAAAAGTTTAAAAATTTTGTGACTTAATATT | 99171 |
| Qy | 544 | AAATACACATAAAAAAGAGGTGTTTAATTAACAAACCATGTTTCGTACTCAAACT--CTA | 601 |
| Db | 99172 | AAACCTTATAATTTATGAAATAATATTAATTAATTTGTATGAAGAAATTTTACTATTAAACA | 99231 |
| Qy | 602 | ATTGTGAAATCTCTATTTTCAGTCACAAAATTTCCAAATTTCCAAATTAAGAAAAATTAACGTA | 661 |
| Db | 99232 | CATGTAATAATTTAATAAAAAATTAATTTAAATTTAAAAATGCAAAAATCATTGTATA | 99291 |
| Qy | 662 | GACGGCTAAGCCACCACCTTAAGGCTAAGTTCGAGAGGTGAAGTAGCGCAGCAAAAATAT | 721 |
| Db | 99292 | TAAAAATGAAATAATATAAAAAATTTATCTTAGTTGTCTATTATTTGTACGGACGCTATTAC | 99351 |
| Qy | 722 | GATGGTTTATTAATATGATTTTTTTTAAATAACTTTTCACATAAATTTCTTTAGGAAACAT | 781 |
| Db | 99352 | GATTATGTTAAATATACAAAATATTAAAAATA--TATAAAAATTATCTCAATTTTAAATAT | 99408 |
| Qy | 782 | ATCATTTAATGGTTTGAAAAACGTCACATTAAGAAAACTAAGAACGATCAGTTGGGAAC | 841 |
| Db | 99409 | ACAAAAATATTTTAAATAAAACATTTATTAATAAATAATATAAAATATTTAAAAATAAAAAA | 99468 |
| Qy | 842 | AAGAGAAAAACACAGCCTTTAAGGCTTCITGATCCTCTAGTTGGAGGTGATTTTCAACG | 901 |
| Db | 99469 | ATATACATATTAAATAAAATCATATAAAAAATACATACAAAAATATAAATAAAACTAAT | 99528 |
| Qy | 902 | CATGATAAACGGAAGCTCATTAGCACATTATTACTTAGATATTTTATTAATATAAACTT | 961 |
| Db | 99529 | AAGAAGTAAAAATTAATAAATATTTATTTAAATTAATTAATAAATAAATAAATTTATTTCA | 99588 |
| Qy | 962 | GAAAAAAATATTTATTTGTAATTTTTTTTAAACAATGATGCATAAATATTTTTTTAAACA | 1021 |
| Db | 99589 | TTAAAAAATTTGMAATATAAAAAAAGAAAAATACAGTAAATTAATAATCCTTAAACA | 99648 |
| Qy | 1022 | CACCAATTTAACCCCTTTAAAAA-----GCATCCTAATAGGAAACGAGAAAGTTTAAAGA | 1074 |
| Db | 99649 | AAAAAATTTTAAACATGAAAAAAAATGACTTTTAAAGTCATAGAAAAAAAACCTTACAAC | 99708 |
| Qy | 1075 | TTCAACCGAAGTGTTTTGGATAATGAATAATGGGTGGGATTTAGATTTGGTAAATGAT--CA | 1133 |
| Db | 99709 | TTCAAAGTGTAAAGTTCAAAGAAAAAATCTGTACGACTTTTAAAAAATGTAAATAAATAAA | 99768 |
| Qy | 1134 | GGGTTAGGATTAATAATTTAAAAATGAAGAGGGAGAATGAATGGTTAGA | 1181 |
| Db | 99769 | AAAAATAAGCTAAAGTTGTAAAAAAAATAAAAAATAAATAAATAAATCGTATA | 99816 |

```

RESULT 10
; US-09-754-853A-3
; Sequence 3, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810) B
; CURRENT APPLICATION NUMBER: US/09/754, 853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 3

```


APPLICANT: Wang, Ming Li
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE OF INVENTION: Soybean Cyst Nematode Resistance
FILE REFERENCE: 38-10(15810)B
CURRENT APPLICATION NUMBER: US/09/754, 853A
CURRENT FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: US 60/174,880
PRIOR FILING DATE: 2000-01-07
NUMBER OF SEQ ID NOS: 1119
SEQ ID NO 3
LENGTH: 335913
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: CDS
LOCATION: (46798)..(48763).(48975)..(49573)
OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-3

Query Match 3.4%; Score 48.8; DB 9; Length 335913;
Best Local Similarity 45.3%; Pred. No. 3;
Matches 354; Conservative 0; Mismatches 407; Indels 21; Gaps 4;

QY 449 AAAAAGAAATTTAATTTTTTTTTTTCATGAAACGCAATTCACCGTTCGAGAAATATGCT 508
DB 183354 AACAAAAAATTTTCAATTTAACTAGTAATTTAAATTTGAATATATGCATTAATATTTT 183295
QY 509 GTCATAAATAAGTAGCTAGTCGACGAAACAAATTAATATCACAATAAAAGAGGTGT 568
DB 183294 TTAATAAATAATTTTACCTCATATAATAATATTTTGAATTTGAATATATCAGCGCTGATGT 183235
QY 569 TAATTACAAAC-----CATGTTTCGTACTACAACTCAATTTGTAAATCTTTATT 618
DB 183234 AAAATGATTTCTTCATCTCTCTTTTATAAGAATAAATCACAATTTTATTTCTCTTTT 183175
QY 619 TCAGTCACAAATTCGAATTTTCCAAATTTAAGAAAAATAACGTAGACGGCTAAGCCACCC 678
DB 183174 TAAGTAAATTTTGTCTATTTTACTTTTAAATGAGATTAATTTAAATATCTTTTCAC 183115
QY 679 ATCTAAGGCTTAAGTTCGAGAGTGAAGTACGCACGAAAAATATGATGTTTATTAATATG 738
DB 183114 TCAACAGGTGTTGCTTAATATCGAGTTTAAATGCATAAATTTAG---AAAAATATG 183058
QY 739 ATTTTTTTAAATACTTTCACATAAATTTCTTTAGGAAACATATCAATTTAATGGTTTGA 798
DB 183057 TTTATTTTAAATTTAAATTTAATATAATTTAATAAAATTAACTTATTTTACTAATAAGTAT 182998
QY 799 AAAAA-----CGTGCACATAAGAAACCTAAAGAACGATGAGTTGGAAACAGAGAAAC 852
DB 182997 GAAATTAATCTTTTCTTTATAGAGAGAGAGGAGTACTTTGTGATTTATTAATAATA 182938
QY 853 ACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTGAATTTCAAACGGCATGATAACG 912
DB 182937 TATTTATATTCACCCCAATGACCTTTTAAATTTTATTTGTCTAATCCTAAAAATAAAT 182878
QY 913 AGAAGCTCATTAGCACATTTATCTAGATATTTATATTAATAAATCTGAAAAAATAT 972
DB 182877 TGAACACATATTTTATTTATTTATAGATTTTAAATGTTATGATCTAATAAGAGATC 182818
QY 973 TTAATTTGAATTTTAAACAATGTATGCATAAATTTATTTTTTAAACACACCAATTTAA 1032
DB 182817 TGGTGTGAATCTATTAAAAAAGTAAGAAAAATTAAGTAAATTTAATAACACCAATGTAC 182758
QY 1033 CCCTTTAAAAAGCATCTTAATAGGAAACGAGCAAGTTAAAGATTCACCGAAGTGTTTGGA 1092
DB 182757 AAGCAACTTAAGCAAGAACAAAGAGTAAGGAAGGGATGTACAAATAGAAAAAGA 182698
QY 1093 TAATGAAAAATGGGTGGGATTTAGAAATTTGGTAAATGAATCAGGGTTAGGATTTAAATTTA 1152
DB 182697 TTACGAGGAAATGAAGAGGAGAAAAAT--GTAGAGAGCAATGTACAAACATAAGAGAA 182640
QY 1153 AATGAGAGGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTTTGTGGGTAGAAA 1212

DB 182639 AAAATATAAGGAAAAAATGTATATGTATACAGAAATAGGTATTAATGGTTTGAATCGAA 182580
QY 1213 TT 1214
DB 182579 TT 182578

RESULT 14
US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. US20020127687A1
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEKI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 3.3%; Score 47.2; DB 10; Length 640681;
Best Local Similarity 48.4%; Pred. No. 7.3;
Matches 199; Conservative 0; Mismatches 208; Indels 4; Gaps 2;

QY 245 TGTGATACCTGCTGTGCACAAAGCATCAAAATTCCTCTCTGGAGTATCTTTATTCGGA 304
DB 580300 TATACCTAGCTGATTTATTAATAAAACAAAGACATTTAAGTTATAAATAAATTTCTTTA 580359
QY 305 AAACCCCAAGATTTATCTTATTCACCTCAGGCTAAATTTGCTGAACTATGCAATGAATA 364
DB 580360 AAGAAAAATCTTTAGTTAATAAATAAACAACAAAAATCAATCATATGTTAGTTCATA 580419
QY 365 CAATTCGCAATATCATGGTTATCTATCTGCTCAAAATTTGAATTTGAGTCCAACCTGA 424
DB 580420 AAAAAATCTAGAAATTTGAT--TTTTTTTATTTAATAAATATGAAATTTTACTTAACT 580477
QY 425 GACTGCAATACGATTTTCTTTTCAAAAAGAAATTTAATAATTTTTCATGAAACGCA 484
DB 580478 GTATATGTTAAAAAATTTACTTAATAAATAAATAAATAAATAAATAAATAAATAA 580537
QY 485 ATTCACCCGTTTCAGAAATATCTGTCATAAA--TAAGTAGCTAGTGCAGAAACAAAT 542
DB 580538 ATAAAAAATTAATTTGTTATAGTAGAATATATCTAAAAAGTTTATTGATTGAAAAAAT 580597
QY 543 TAATATCACATAAAAAAGAGGTTGTTAATTAACAACCATGTTTCGTACTCAACTCTAA 602
DB 580598 AATATTAAGAAATATAAATAATTTTACATTGAAAAACAGATCAGAAAAATATATTA 580657
QY 603 TTTGTAAATTTCTTTCTAGTCACAAAATTTCAATTTCCAAATTTAAGAAAAA 653
DB 580658 TTTTATCTAAACCTAATTTGGAATAATTAATAAATAAATAAATAAATAAATAA 580708

RESULT 15
US-09-352-11234/c
Sequence 11234, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Ningbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11234
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 48-LIB3058-052-Q1-K1-D8
US-09-960-352-11234

Query Match 3.2%; Score 46.6; DB 10; Length 419;
Best Local Similarity 48.3%; Pred. No. 1.6;
Matches 130; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 892 TTTTCARACGCATGATAAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATATAA 951
DB 295 TTTTAAAAAATTTTAAAAATTTTAAATTTTAAATTTTAAAAATATAATATAAAAAATTTT 236
QY 952 TTATAAACTTCAAAAAATAATTTTGAATTTTAAACAATGTATGCATAAATTTT 1011
DB 235 TTTTATATTTAATAAAATTTATAAATTTTAACTTTTAAAAATTTTATTTATTTT 176
QY 1012 TTTAAAAACACACCAATTTAACCCCTTTAAAAAGCATCTAATAGGAAACGAGGAAGTTAA 1071
DB 175 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 116
QY 1072 AGATTCACCGAAGTGTTTGGATAATGAAAAATGGGGTGGGATTGAAATGGTAAATGAAT 1131
DB 115 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 56
QY 1132 CAGGCTTAGGATTAAATATTAAATGAAA 1160
DB 55 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 27

Search completed: March 15, 2003, 20:58:48
Job time : 1764.39 secs